

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 16, 2001, 16:12:49 : Search time 17.94 Seconds  
(without alignments)  
638.513 Million cell updates/sec

Title: US-09-325-019-2  
Perfect score: 1937  
Sequence: 1 DFTPAFLDTSRRPQFCCKMP.....NPNDIFADLESYPDESEIAN 335

Scoring table:  
BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 268485 seqs, 34193795 residues  
Total number of hits satisfying chosen parameters: 268485

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

A.Geneseq.-36:\*

- 1: /SIDSL/gcgdata/geneseq/geneseq/AA1980.DAT:\*
- 2: /SIDSL/gcgdata/geneseq/geneseq/AA1981.DAT:\*
- 3: /SIDSL/gcgdata/geneseq/geneseq/AA1982.DAT:\*
- 4: /SIDSL/gcgdata/geneseq/geneseq/AA1983.DAT:\*
- 5: /SIDSL/gcgdata/geneseq/geneseq/AA1984.DAT:\*
- 6: /SIDSL/gcgdata/geneseq/geneseq/AA1985.DAT:\*
- 7: /SIDSL/gcgdata/geneseq/geneseq/AA1986.DAT:\*
- 8: /SIDSL/gcgdata/geneseq/geneseq/AA1987.DAT:\*
- 9: /SIDSL/gcgdata/geneseq/geneseq/AA1988.DAT:\*
- 10: /SIDSL/gcgdata/geneseq/geneseq/AA1989.DAT:\*
- 11: /SIDSL/gcgdata/geneseq/geneseq/AA1990.DAT:\*
- 12: /SIDSL/gcgdata/geneseq/geneseq/AA1991.DAT:\*
- 13: /SIDSL/gcgdata/geneseq/geneseq/AA1992.DAT:\*
- 14: /SIDSL/gcgdata/geneseq/geneseq/AA1993.DAT:\*
- 15: /SIDSL/gcgdata/geneseq/geneseq/AA1994.DAT:\*
- 16: /SIDSL/gcgdata/geneseq/geneseq/AA1995.DAT:\*
- 17: /SIDSL/gcgdata/geneseq/geneseq/AA1996.DAT:\*
- 18: /SIDSL/gcgdata/geneseq/geneseq/AA1997.DAT:\*
- 19: /SIDSL/gcgdata/geneseq/geneseq/AA1998.DAT:\*
- 20: /SIDSL/gcgdata/geneseq/geneseq/AA1999.DAT:\*
- 21: /SIDSL/gcgdata/geneseq/geneseq/AA2000.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1937	100.0	335	21	Y59247 Human connective t
2	1937	100.0	345	20	Y17640 Human putative mat
3	1937	100.0	367	20	Y17641 Human WISP-1 prote
4	1936	99.9	345	20	Y17642 Human WISP-1 varia
5	1936	99.9	367	20	Y17643 Human WISP-1 varia
6	1934	99.8	345	20	Y17643 Human WISP-1 varia
7	1934	99.8	367	20	Y17645 Human WISP-1 varia
8	1933	99.8	345	20	Y17652 Human WISP-1 varia
9	1933	99.8	367	20	Y17653 Human WISP-1 varia
10	1675	86.5	345	20	Y17646 Mouse putative mat
11	1675	86.5	367	20	Y17647 Mouse WISP-1 prote
12	869.5	44.9	349	16	R79964 Connective tissue

RESULT	ALIGNMENTS
13	869.5 44.9 349 18 W09089 Human connective t
14	869.5 44.9 349 18 W1302 Connective tissue
15	869.5 44.9 349 19 W62084 Human connective t
16	869.5 44.9 349 20 Y19361 Human connective t
17	869.5 44.9 349 20 W81425 Connective tissue
18	869.5 44.9 349 21 Y44755 Human connective t
19	868.5 44.8 347 18 W12694 Connective tissue
20	846 43.7 339 20 Y17654 Human putative mat
21	846 43.7 372 20 Y17655 Human WISP-3 prote
22	844 43.6 339 20 Y17656 Human putative mat
23	844 43.6 354 20 Y34190 Human connective t
24	844 43.6 354 21 Y17657 Human WISP-3 prote
25	844 43.6 354 21 Y81438 Human growth facto
26	835.5 43.1 347 20 Y24379 Rat connective tis
27	835.5 43.1 348 13 R25566 Beta-IG-M2. Mus m
28	831.5 42.9 348 18 W35731 Murine FISP12. Mu
29	831.5 42.9 348 21 Y44756 Bovine connective t
30	793.5 41.0 349 21 Y44754 Mouse connective t
31	772.5 39.9 379 13 R25565 Beta-IG-M1. Mus m
32	764.5 39.5 381 18 W35957 Human monocyte mat
33	760.5 39.3 381 18 W35730 Human cysteine ric
34	743.5 38.4 351 14 R31599 Chicken nov protei
35	705.5 36.4 375 17 R90919 Connective tissue
36	705.5 36.4 375 20 Y31620 Human CTGF-2. Hom
37	498.5 25.7 227 20 Y27440 Rat HICP mature po
38	498.5 25.7 250 20 Y27434 Mouse putative mat
39	498 25.7 228 20 Y17650 Mouse WISP-2 prote
40	498 25.7 229 20 Y17701 Mouse WISP-2 prote
41	498 25.7 230 20 Y17700 Mouse WISP-2 prote
42	498 25.7 231 20 Y17699 Mouse WISP-2 prote
43	498 25.7 232 20 Y17698 Mouse WISP-2 prote
44	498 25.7 233 20 Y17697 Mouse WISP-2 prote
45	498 25.7 234 20 Y17696 Mouse WISP-2 prote

RESULT	1	Location/Qualifiers
Y59247	Y59247 standard; Protein: 335 AA.	
XX		
AC	Y59247;	
XX		
DT	11-APR-2000 (first entry)	
DE		
XX		
XX	Human connective tissue growth factor-4 (CTGF-4).	
KW	Connective tissue growth factor-4; CTGF-4; human; immune system disorder;	
KW	hematopoietic disorder; autoimmune disorder; diabetes mellitus; asthma;	
KW	respiratory disorder; inflammation; hyperproliferative disorder;	
KW	infection; central nervous system disease; Alzheimer's disease; AIDS;	
KW	food additive.	
XX		
OS	Homo sapiens.	
XX		
FH	Key	
FT	Domain	15..84
FT	Domain	/note- "IGF binding domain"
FT	Domain	28..36
FT	Domain	/note- "conserved domain CD-I"
FT	Domain	39..55
FT	Domain	/note- "conserved domain CD-II"
FT	Domain	54
FT	Domain	/note- "potential N-glycosylation site"
FT	Domain	61..70
FT	Domain	/note- "conserved domain CD-III"
FT	Domain	89..154
FT	Domain	/note- "Von-Willebrand factor type C repeat fragment"
FT	Domain	101..121
FT	Domain	/note- "conserved domain CD-IV"
FT	Domain	111
FT	Modified-site	/note- "potential N-glycosylation site"

FT	Domain	144..154	/note="conserved domain CD-V"
FT	Domain	184..228	/note="sulfated glycoconjugate binding motif"
FT	Domain	194..213	/note="conserved domain CD-VI"
FT	Domain	216..227	/note="conserved domain CD-VII"
FT	Domain	236..241	/note="conserved domain CD-VIII"
FT	Domain	241..316	/note="C-terminal dimerisation and receptor-binding domain"
FT	Modified-site	252	
FT	Domain	253..260	/note="potential N-glycosylation site"
FT	Domain	264..280	/note="conserved domain CD-IX"
FT	Domain	290..295	/note="conserved domain CD-X"
FT	Domain	311	/note="conserved domain CD-XI"
FT	Modified-site		/note="potential N-glycosylation site"
PN	WC0962927-A1.		
XX	09-DEC-1999.		
PD	03-JUN-1999;	99WC-US12150.	
PF	05-JUN-1998;	98US-0088320.	
PR	(HUMA-) HUMAN GENOME SCI INC.		
PA	Ruben SM, Young PE;		
PI	WPI: 2000-147042/13.		
PI	DR N-PSDB: Z58613.		
XX	New isolated connective tissue growth factor-4, used for treating e.g. cancers -		
PS	Claim 11, Fig 1A-E; 196pp; English.		
XX	The invention provides an isolated human connective tissue growth factor -4 (CTGF-4) polypeptide. The CTGF-4 cDNA is deposited under ATCC No. 209816. The CTGF-4 protein can be expressed by standard recombinant methodology. The polypeptides can be used for preventing, treating or ameliorating a medical condition. They may be useful in treating deficiencies or disorders of the immune system, by activating or inhibiting the proliferation, differentiation, or mobilization (chemotaxis) of immune cells, treating or detecting deficiencies or disorders of hematopoietic cells (e.g. blood protein disorders, ataxia telangiectasia, HIV infection, Di-George syndrome, anemia or thrombolytic activity (clot formation) (e.g. blood coagulation disorders, blood platelet disorders, or wounds resulting from trauma, or surgery), in treating or detecting autoimmune disorders (e.g. Addison's disease, rheumatoid arthritis, allergic encephalomyelitis, Goodpasture's syndrome, multiple sclerosis, purpura, Reiter's disease, Guillain-Barre syndrome, systemic lupus erythematosus, insulin dependent diabetes mellitus or autoimmune inflammatory eye disease), treating asthma (particularly allergic asthma) or other respiratory problems (e.g. anaphylaxis, hyperreactivity, to an antigenic molecule or blood group incompatibility, to treat and/or prevent organ rejection or graft-versus-host disease (GVHD), to modulate inflammation (septic shock, sepsis, arthritis, nephritis, cytokine or chemokine induced lung injury, inflammatory bowel disease, Crohn's disease, or resulting from over production of cytokines), to treat hyperproliferative disorders, including neoplasms in the abdomen, bone, breast, digestive system, liver, pancreas, peritoneum, endocrine glands, eye, head and neck, nervous (central and peripheral), lymphatic system, pelvic, skin, soft tissue, spleen, choroid and urogenital, hypergammaglobulinemia,		

CC	lymphoproliferative disorders, Waldenstrom's macroglobulinemia, sarcoidosis), to treat or detect infectious agents, e.g. viruses (e.g. arthritits, bronchiolitis, encephalitis, eye infections, chronic fatigue syndrome, hepatitis, meningitis, AIDS, pneumonia, chickenpox, measles, mumps, parainfluenza, rabies, the common cold, polio, leukemia, rubella, sexually transmitted diseases, or skin diseases) bacterial or fungal agents (e.g. bacteremia, endocarditis, eye infections, gingivitis, opportunistic infections, respiratory tract infections, Lyme disease, cat-scratch disease, paratyphoid fever, food poisoning, pneumonia, gonorrhea and sexually transmitted diseases, meningitis, tuberculosis, lupus, gangrene, tetanus, rheumatic fever, urinary tract infections, wound infections), parasitic agents (e.g. scabies, dysentery, liver disease, malaria, toxoplasmosis), to differentiate, proliferate and attract cells, leading to the regeneration of tissues (e.g. repair, replace or protect tissue in wounds, burns, incisions or ulcers, osteoporosis, osteoarthritis, periodontal disease, liver failure, surgery, cosmetic plastic surgery, reperfusion injury) to proliferate and differentiate nerve cells (e.g. spinal cord disorders, head trauma, cerebrovascular disease and stroke), localized neuropathies and central nervous system diseases (e.g. Alzheimer's disease, Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis, and Shy-Drager syndrome). They may also increase or decrease the differentiation or proliferation of embryonic stem cells and hematopoietic lineage, may be used to modulate mammalian characteristics such as body height, weight, hair color, eye color, skin, percentage of adipose tissue, pigmentation, size, and shape, to modulate mammalian metabolism affecting catabolism, anabolism, processing, utilization and storage of energy, to change a mammal's mental state or physical state by influencing biorhythms, cardiac rhythms, circadian rhythms, depression (including depressive disorders), tendency for violence, tolerance for pain, reproductive capabilities, hormonal or endocrine levels, appetite, libido, memory, stress, or other cognitive qualities, as a food additive or preservative, such as to increase or decrease storage capabilities, fat content, lipid, protein, carbohydrate, vitamins, minerals, cofactors or other nutritional components. Mutations in the PNS or the presence or amount of expression or activity of the polypeptides can be used for diagnosing a pathological condition or a susceptibility to a pathological condition. The CTRF-4 polypeptides can also be used for identifying binding partners. The CC products can also be used for producing transgenic animals. The present sequence represents the CTRF-4 polypeptide.	CC
CC	XX	CC
CC	Sequence	335 AA;
CC	Query Match	100.0%; Score 1937; DB 21; Length 335;
CC	Best Local Similarity	100.0%; Pred. No. 4e-138;
CC	Matches 335; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
CC	1	DFTPALEDTSSRPFCKPCBCPPSPRCPLGVSLTIDGCECCMKCAQOAGDNCTEAT 60
CC	1	dftpaledtssrpfckpcbcppsprcplgvslltdgceccmkcagqigdncleaai 60
CC	61	CDPRHGATCDVSGDRBRRAIVGCAQVVGVLDSVRNNQSSPPNCKNYNCTCIDXAVG 120
CC	61	cdprhgatcdvsgdrbrraivgcaqvvgvlldsvrnnqssppncknyntcidxavg 120
CC	121	CTPICLRPRRLMCPHRRRYSIPGHCEQWVCEEDAKRPRKTPARPDGADAVEGVAM 180
CC	121	ctpiclrprrlmcprrrrysipghceqvwvceddkrrprktpardgadavegvam 180
CC	181	HRNCIAVTSPPSPCSTGCLGVSTRISVNNNAQWDEQSRICNLPRCDVDITLTKAGK 240
CC	181	hrnciavtspspcstgclgvstrisvnnnaqwdeqsrilcnlprcdvdtltlkagk 240
CC	241	CLAYVPPASNNPTLAGISTRVQPKKCGVCMNRCCIPKSKRIDVSFOCPDGLGSR 300
CC	241	clayvppasnmftlagistrvqpkkcygvcmdnrccipyksklidvsfgcpdglgsr 300
CC	301	QVLWLNACFNILSCRNPNDIFADLESYDFSEIAN 335
CC	301	qvlwnacfnilscrnpndifadlesydfseian 335

```

RESULT 2
Y17640
ID Y17640 standard; Protein; 345 AA.
XX
AC Y17640;
XX
DT 06-AUG-1999 (first entry)
XX
DE Human putative mature WISP-1 protein SEQ ID NO:3.
XX
KW WNT-1 induced secreted protein; WISP-1; WISP-2; WISP-3; CTGF; tumour;
KW connective tissue growth factor; cancer; melanoma; arteriosclerosis;
KW leukemia; lymphoid malignancy; haematopoiesis-related disorder;
KW tissue-growth disorder; skin disorder; desmoplasia; fibrotic lesion;
KW kidney disorder; bone-related disorder; osteoporosis; trauma; burn;
KW connective tissue disorder; catabolic state; inflammation;
KW testicular-related disorder; angiogenesis; immunological disorder.
XX
OS Homo sapiens.
XX
PN MO9921998-A1.
XX
PD 06-MAY-1999.
XX
PF 29-OCT-1998; 98MO-US22991.
XX
PR 14-APR-1998; 98US-0081695.
PR 29-OCT-1997; 97US-0063704.
PR 03-FEB-1998; 98US-0073612.
XX
XX
PA (GETH ) GENENTECH INC.
XX
PI Botstein DA, Cohen RU, Goddard A, Gurney AL, Hillan K;
PI Lawrence DA, Levine AJ, Pennica D, Roy MA, Wood WI;
XX
DR WPI: 1999-337420/28.
XX
PT New isolated Wnt-1 induced secreted polypeptides, WISP-1, 2 and 3
XX
PS Claim 1; Page 162-163; 284pp; English.
XX
XX
CC The present invention describes Wnt-1 induced secreted polypeptides,
CC WISP-1, 2 and 3. The novel WISP polypeptides, designated WISP-1, WISP-2
CC and WISP-3 have homology to connective tissue growth factor (CTGF).
CC Products from the present invention can be used to treat WISP-related
CC disorders such as breast, ovarian, and colon cancer or melanoma. The
CC products can be used to treat arteriosclerosis. The products can also be
CC used to treat other diseases e.g. benign and malignant tumours,
CC leukaemia and lymphoid malignancies, neuronal, glial, astrocytal,
CC hypothalamic and other glandular, macrophagal, epithelial, stromal, and
CC blastocoelic disorders, haematopoesis-related disorders, tissue-growth
CC disorders, skin disorders, desmoplasia, fibrotic lesions, kidney
CC disorders, bone-related disorders such as osteoporosis, trauma such as
CC burns, incisions, and other wounds, connective tissue disorders,
CC catabolic states, testicular-related disorders, and inflammatory,
CC angiogenic and immunologic disorders including arteriosclerosis. The
CC products can also be used for detection and diagnosis especially of
CC individuals with neoplastic cell growth or proliferation. The products
CC can be used in the production of transgenic or knock-out animals.
CC Antihodles can be used to induce death in WISP-1, 2 or 3 overexpressing
CC cells.
XX
SQ Sequence 345 AA;

```

```

Query Match 100.0%; Score 1937; DB 20; Length 345;
Best Local Similarity 100.0%; Pred. No. 4, 2e-138;
Matches 335; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 DFTPPAPLEDTSSRPQCKWPCPCPPSPPLGVSILITDGCCKMKCAQQLGDNCTEAAI 60
DB 11 dftppapleDTSSRPQCKWPCPCPPSPPLGVSILITDGCCKMKCAQQLGDNCTEAAI 70

```

```

QY 61 CDPHRLGYCDYSGDRPRVAYIGCAQYVVGCVLDBGVRYNNGQSRQPNCKYKNCITDCAVG 120
DB 71 CDPHRLGYCDYSGDRPRVAYIGCAQYVVGCVLDBGVRYNNGQSRQPNCKYKNCITDCAVG 130
QY 121 CTPCLRLVRPRLMPCPHRRVRSIRPHCCEDQWVCEBDARPRKTPAPRDTGAFDAVEAEAW 180
DB 131 CTPCLRLVRPRLMPCPHRRVRSIRPHCCEDQWVCEBDARPRKTPAPRDTGAFDAVEAEAW 190
QY 181 HRCNIAVTSFMSPCSTSCGLGVSTRISNVNAQCPBEDSRICNLRPCDVIDHTLTKAGKK 240
DB 191 HRCNIAVTSFMSPCSTSCGLGVSTRISNVNAQCPBEDSRICNLRPCDVIDHTLTKAGKK 250
QY 241 CLAYQPEASMNFLACISIRSVQPKYCGVCMNRCCIPKSKTIDVSPCCPGICFSR 300
DB 251 CLAYQPEASMNFLACISIRSVQPKYCGVCMNRCCIPKSKTIDVSPCCPGICFSR 310
QY 301 QVLMINACFCNLSCRNPNDIFADLESYPDESEIAN 335
DB 311 QVLMINACFCNLSCRNPNDIFADLESYPDESEIAN 345

```

```

RESULT 3
Y17641
ID Y17641 standard; Protein; 367 AA.
XX
AC Y17641;
XX
DT 06-AUG-1999 (first entry)
XX
DE Human WISP-1 protein SEQ ID NO:4.
XX
KW WNT-1 induced secreted protein; WISP-1; WISP-2; WISP-3; CTGF; tumour;
KW connective tissue growth factor; cancer; melanoma; arteriosclerosis;
KW leukemia; lymphoid malignancy; haematopoesis-related disorder;
KW tissue-growth disorder; skin disorder; desmoplasia; fibrotic lesion;
KW kidney disorder; bone-related disorder; osteoporosis; trauma; burn;
KW connective tissue disorder; catabolic state; inflammation;
KW testicular-related disorder; angiogenesis; immunological disorder.
XX
OS Homo sapiens.
XX
PN MO9921998-A1.
XX
PD 06-MAY-1999.
XX
PF 29-OCT-1998; 98MO-US22991.
XX
PR 14-APR-1998; 98US-0081695.
PR 29-OCT-1997; 97US-0063704.
PR 03-FEB-1998; 98US-0073612.
XX
XX
PA (GETH ) GENENTECH INC.
XX
PI Botstein DA, Cohen RU, Goddard A, Gurney AL, Hillan K;
PI Lawrence DA, Levine AJ, Pennica D, Roy MA, Wood WI;
XX
DR WPI: 1999-337420/28.
XX
PT New isolated Wnt-1 induced secreted polypeptides, WISP-1, 2 and 3
XX
PS Claim 4; Page 163-164; 284pp; English.
XX
XX
CC The present invention describes Wnt-1 induced secreted polypeptides,
CC WISP-1, 2 and 3. The novel WISP polypeptides, designated WISP-1, WISP-2
CC and WISP-3 have homology to connective tissue growth factor (CTGF).
CC Products from the present invention can be used to treat WISP-related
CC disorders such as breast, ovarian, and colon cancer or melanoma. The
CC products can be used to treat arteriosclerosis. The products can also be
CC used to treat other diseases e.g. benign and malignant tumours,
CC leukaemia and lymphoid malignancies, neuronal, glial, astrocytal,
CC hypothalamic and other glandular, macrophagal, epithelial, stromal, and
CC blastocoelic disorders, haematopoesis-related disorders, tissue-growth

```

CC disorders, skin disorders, desmoplasia, fibrotic lesions, kidney  
 CC disorders, bone-related disorders such as osteoporosis, trauma such as  
 CC burns, incisions, and other wounds, connective tissue disorders,  
 CC catabolic states, testicular-related disorders, and inflammatory,  
 CC angogenic and immunologic disorders including arteriosclerosis. The  
 CC products can also be used for detection and diagnosis especially of  
 CC individuals with neoplastic cell growth or proliferation. The products  
 CC can be used in the production of transgenic or knock-out animals.  
 CC Antibodies can be used to induce death in WISP-1, 2 or 3 overexpressing  
 CC cells.

XX Sequence 367 AA;

Query Match 100.0%; Score 1937; DB 20; Length 367;  
 Best Local Similarity 100.0%; Pred. No. 4,4e-138;  
 Matches 335; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DFTPALEDTSSRPOCKWPCPCPPSPRCPLGVSILITDGCCKMCAQQLDNCTEAI 60  
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 DB 33 dftppledtsrpfckwpcpcppsprcplgvsilittgcccckmcaqqlgdncteaal 92  
 QY 61 CDPHRLCYDYSGBDRPRAIGVCAOVYGVLDGVRVNNGSGFQPNCKYNTCTIDGAVG 120  
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 DB 93 cdpnrglycdysgdrpryagvcaovgyvldgyrvnnngsfqpnckynctcidgavg 152  
 QY 121 CTPPLCLRAPPRLMCPHPRRVSIPGHCCQWVCEDDAKRPRTAPRDGAFVGEVAM 180  
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 DB 153 ctpplclrapprlmcprrvsipghccqvwvceddakrprrtaprdtgaftavgeveam 212  
 QY 181 HRNCIAVTPSPGSCSTGSLGVSTRISNVNACWPEQESRLNLRCPDVIHTLTKAGKK 240  
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 DB 213 hnciavtspgscstgslgvstrisnvnacwpeqesrlnlrpdvhlthlkagkk 272  
 QY 241 CLAVTQPEASMTFLAGCISTRSYQPKYGVCMNDNRCIPYKSKTIDVSPQDGLGFSR 300  
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 DB 273 clavyqpeasmftlagcistrsyqpkycgvcmndnrcipysktdvsgfdqpdglgfsr 332  
 QY 301 QVLMINACFCNLSCRNPNDIFADLESYPDFSEIAN 335  
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 DB 333 qvlinacfcnlscrnpndifadlesypdfseian 367

RESULT 4

Y17642 standard; Protein: 345 AA.

XX Y17642;

DT 06-AUG-1999 (first entry)

DE Human WISP-1 variant protein SEQ ID NO:5.

XX WNT-1 induced secreted protein; WISP-1; WISP-2; WISP-3; CTGF; tumour;  
 KW connective tissue growth factor; cancer; melanoma; arteriosclerosis;  
 KW leukaemia; lymphoid malignancy; hematopoiesis-related disorder;  
 KW tissue-growth disorder; skin disorder; desmoplasia; fibrotic lesion;  
 KW kidney disorder; bone-related disorder; osteoporosis; trauma; burn;  
 KW connective tissue disorder; catabolic state; inflammation;  
 KW testicular-related disorder; angiogenesis; immunological disorder.

XX Synthetic.

OS Homo sapiens.

XX MO9921998-A1.

XX 06-MAY-1999.

XX 29-OCT-1998; 98WO-US22991.

XX 14-APR-1998; 98US-0081695.

XX 29-OCT-1997; 97US-0063704.

XX 03-FEB-1998; 98US-0073612.

XX (GENTH ) GENENTECH INC.  
 PA Botstein DA, Cohen RL, Goddard A, Gurney AL, Hillan K;  
 PI Lawrence DA, Levine AJ, Pennica D, Roy MA, Wood WT;  
 DR WPT: 1999-337420/28.

XX New isolated Wnt-1 induced secreted polypeptides, WISP-1, 2 and 3  
 PS Claim 5: Page 164-165; 284pp; English.

CC The present invention describes Wnt-1 induced secreted polypeptides,  
 CC WISP-1, 2 and 3. The novel WISP polypeptides, designated WISP-1, WISP-2  
 CC and WISP-3 have homology to connective tissue growth factor (CTGF).  
 CC Products from the present invention can be used to treat WISP-related  
 CC disorders such as breast, ovarian, and colon cancer or melanoma. The  
 CC products can be used to treat arteriosclerosis. The products can also be  
 CC used to treat other diseases e.g. benign and malignant tumours,  
 CC leukaemia and lymphoid malignancies, neuronal, glial, astrocytal,  
 CC hypothalamic and other glandular, macropagal, epithelial, stromal, and  
 CC blastocellular disorders, haematopoiesis-related disorders, tissue-growth  
 CC disorders, skin disorders, desmoplasia, fibrotic lesions, kidney  
 CC disorders, bone-related disorders such as osteoporosis, trauma such as  
 CC burns, incisions, and other wounds, connective tissue disorders,  
 CC catabolic states, testicular-related disorders, and inflammatory,  
 CC angogenic and immunologic disorders including arteriosclerosis. The  
 CC products can also be used for detection and diagnosis especially of  
 CC individuals with neoplastic cell growth or proliferation. The products  
 CC can be used in the production of transgenic or knock-out animals.  
 CC Antibodies can be used to induce death in WISP-1, 2 or 3 overexpressing  
 CC cells.

SO Sequence 345 AA;

Query Match 99.9%; Score 1936; DB 20; Length 345;  
 Best Local Similarity 99.7%; Pred. No. 4,9e-138;  
 Matches 334; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 DFTPALEDTSSRPOCKWPCPCPPSPRCPLGVSILITDGCCKMCAQQLDNCTEAI 60  
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 DB 11 dftppledtsrpfckwpcpcppsprcplgvsilittgcccckmcaqqlgdncteaal 70  
 QY 61 CDPHRLCYDYSGBDRPRAIGVCAOVYGVLDGVRVNNGSGFQPNCKYNTCTIDGAVG 120  
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 DB 71 cdpnrglycdysgdrpryagvcaovgyvldgyrvnnngsfqpnckynctcidgavg 130  
 QY 121 CTPPLCLRAPPRLMCPHPRRVSIPGHCCQWVCEDDAKRPRTAPRDGAFVGEVAM 180  
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 DB 131 ctpplclrapprlmcprrvsipghccqvwvceddakrprrtaprdtgaftavgeveam 190  
 QY 181 HRNCIAVTPSPGSCSTGSLGVSTRISNVNACWPEQESRLNLRCPDVIHTLTKAGKK 240  
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 DB 191 hnciavtspgscstgslgvstrisnvnacwpeqesrlnlrpdvhlthlkagkk 250  
 QY 241 CLAVTQPEASMTFLAGCISTRSYQPKYGVCMNDNRCIPYKSKTIDVSPQDGLGFSR 300  
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 DB 251 clavyqpeasmftlagcistrsyqpkycgvcmndnrcipysktdvsgfdqpdglgfsr 310  
 QY 301 QVLMINACFCNLSCRNPNDIFADLESYPDFSEIAN 335  
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 DB 311 qvlinacfcnlscrnpndifadlesypdfseian 345

RESULT 5

Y17644 standard; Protein: 367 AA.

XX Y17644;

DT 06-AUG-1999 (first entry)

DE Human WISP-1 variant protein SEQ ID NO:7.

XX  
XX WNT-1 induced secreted protein; WISP-1; WISP-2; WISP-3; CTGF; tumour;  
KW connective tissue growth factor; cancer; melanoma; arteriosclerosis;  
KW leukemia; lymphoid malignancy; haematopoiesis-related disorder;  
KW tissue-growth disorder; skin disorder; desmoplasia; fibrotic lesion;  
KW kidney disorder; bone-related disorder; osteoporosis; trauma; burn;  
KW connective tissue disorder; catabolic state; inflammation;  
KW testicular-related disorder; angiogenesis; immunological disorder.  
XX  
XX Synthetic.  
OS Homo sapiens.  
XX  
XX MO9921998-A1.  
XX  
XX 06-MAY-1999.  
XX  
XX 29-OCT-1998; 98MO-US22991.  
XX  
XX 14-APR-1998; 98US-0081695.  
PR 29-OCT-1997; 97US-0063704.  
PR 03-FEB-1998; 98US-0073612.  
XX  
XX (GETH ) GENENTECH INC.  
XX  
PI Botstein DA, Cohen RL, Goddard A, Gurney AL, Hillan K;  
PI Lawrence DA, Levine AJ, Pennica D, Roy MA, Wood WI;  
XX  
XX WPI; 1999-337420/28.  
XX  
XX New isolated Wnt-1 induced secreted polypeptides, WISP-1, 2 and 3  
XX  
XX Claim 6; Page 167-168; 284pp; English.  
XX  
XX The present invention describes Wnt-1 induced secreted polypeptides,  
CC WISP-1, 2 and 3. The novel WISP polypeptides, designated WISP-1, WISP-2  
CC and WISP-3 have homology to connective tissue growth factor (CTGF).  
CC Products from the present invention can be used to treat WISP-related  
CC disorders such as breast, ovarian, and colon cancer or melanoma. The  
CC products can be used to treat arteriosclerosis. The products can also be  
CC used to treat other diseases e.g. benign and malignant tumours,  
CC leukaemia and lymphoid malignancies, neuronal and malignant tumours,  
CC hypochromic and other glandular, macrophagal, epithelial, stromal, and  
CC blastocoele disorders, haematopoiesis-related disorders, tissue-growth  
CC disorders, skin disorders, desmoplasia, fibrotic lesions, kidney  
CC disorders, bone-related disorders such as osteoporosis, trauma such as  
CC burns, incisions, and other wounds, connective tissue disorders,  
CC catabolic states, testicular-related disorders, and inflammatory,  
CC angiogenic and immunologic disorders including arteriosclerosis. The  
CC products can also be used for detection and diagnosis especially of  
CC individuals with neoplastic cell growth or proliferation. The products  
CC can be used in the production of transgenic or knock-out animals.  
CC Antibodies can be used to induce death in WISP-1, 2 or 3 overexpressing  
CC cells.  
XX  
XX Sequence 367 AA;  
XX

Query Match 99.9%; Score 1936; DB 20; Length 367;  
Best Local Similarity 99.7%; Pred. No. 5, 3e-138;  
Matches 334; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 DTPPALEDTSSRPQFCWPCPCPPSPRCPLGVSLITDGCCECKMCAQOLGDNCTEAAI 60  
DB 33 dftppaledtssrpqfcwpcpcppsprrcpilgysliltgdcceckmcaqolgdncetaai 92  
QY 61 CDHHRGIXCDYSGDRPRXAYAGCAQVYGVGVYNNQSGQPRKXNCTCIDANG 120  
DB 93 cdprrgixcdysgdrprxayagcaqvvygvvnnqsgqprkxnyctcidang 152  
QY 121 CTFPLCLTRVRRPRLMCPHRRVRSIPGHCCCEQWVCEDDAKRRPKTPAPRDTGAFDAVGEVAAW 180  
DB 153 ctfplcltrvrprrlmcphrrvrsipghccceqvwceddakrrpktrpdtgafdavgeveaw 212

QY 181 HRNCIATSPWSPCSTSCGLGVSTRISNVNNAQCWPBESRLCNRPDVIDHTLIRAGKK 240  
DB 213 hrnciatspwspcstscglgvstrisnvnaqcwpbesrlcnrpdvidhtliragkk 272  
QY 241 CLAVYQEAAMNFIACISITRSYQPKYCGVCMNRCCIPKSKTIDVSPCCPGLGFSR 300  
DB 273 clavyqeamnfiacisitrsyqpkycgvcmnrccipksktidvspccpglgfsr 332  
QY 301 QVIMINACFCNLSGRNPNIDFADLESYPDFSEIAN 335  
DB 333 qviminacfcnlsgnrpnidfadlesypdfseian 367

RESULT 6  
ID Y17643  
XX Y17643 standard; Protein; 345 AA.  
XX  
AC Y17643;  
XX  
XX 06-AUG-1999 (first entry)  
XX  
XX Human WISP-1 variant protein SEQ ID NO:6.  
XX  
XX WNT-1 induced secreted protein; WISP-1; WISP-2; WISP-3; CTGF; tumour;  
KW connective tissue growth factor; cancer; melanoma; arteriosclerosis;  
KW leukemia; lymphoid malignancy; haematopoiesis-related disorder;  
KW tissue-growth disorder; skin disorder; desmoplasia; fibrotic lesion;  
KW kidney disorder; bone-related disorder; osteoporosis; trauma; burn;  
KW connective tissue disorder; catabolic state; inflammation;  
KW testicular-related disorder; angiogenesis; immunological disorder.  
XX  
XX Synthetic.  
OS Homo sapiens.  
XX  
XX MO9921998-A1.  
XX  
XX 06-MAY-1999.  
XX  
XX 29-OCT-1998; 98MO-US22991.  
XX  
XX 14-APR-1998; 98US-0081695.  
PR 29-OCT-1997; 97US-0063704.  
PR 03-FEB-1998; 98US-0073612.  
XX  
XX (GETH ) GENENTECH INC.  
XX  
PI Botstein DA, Cohen RL, Goddard A, Gurney AL, Hillan K;  
PI Lawrence DA, Levine AJ, Pennica D, Roy MA, Wood WI;  
XX  
XX WPI; 1999-337420/28.  
XX  
XX New isolated Wnt-1 induced secreted polypeptides, WISP-1, 2 and 3  
XX  
XX Claim 5; Page 166-167; 284pp; English.  
XX  
XX The present invention describes Wnt-1 induced secreted polypeptides,  
CC WISP-1, 2 and 3. The novel WISP polypeptides, designated WISP-1, WISP-2  
CC and WISP-3 have homology to connective tissue growth factor (CTGF).  
CC Products from the present invention can be used to treat WISP-related  
CC disorders such as breast, ovarian, and colon cancer or melanoma. The  
CC products can be used to treat arteriosclerosis. The products can also be  
CC used to treat other diseases e.g. benign and malignant tumours,  
CC leukaemia and lymphoid malignancies, neuronal and malignant tumours,  
CC hypochromic and other glandular, macrophagal, epithelial, stromal, and  
CC blastocoele disorders, haematopoiesis-related disorders, tissue-growth  
CC disorders, skin disorders, desmoplasia, fibrotic lesions, kidney  
CC disorders, bone-related disorders such as osteoporosis, trauma such as  
CC burns, incisions, and other wounds, connective tissue disorders,  
CC catabolic states, testicular-related disorders, and inflammatory,  
CC angiogenic and immunologic disorders including arteriosclerosis. The  
CC products can also be used for detection and diagnosis especially of  
CC individuals with neoplastic cell growth or proliferation. The products

CC can be used in the production of transgenic or knock-out animals.  
CC Antibodies can be used to induce death in WISP-1, 2 or 3 overexpressing  
CC cells.

XX Sequence 345 AA;

Query Match 99.8%; Score 1934; DB 20; Length 345;  
Best Local Similarity 99.7%; Pred. No. 7e-138;  
Matches 334; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 DFTPAPLEDTSSRPQFCWMPCECPSPRCPLGVSILITDGCBCCKMAQQLGDNCTEAI 60  
DB 11 dftppledtssrpfqfckwpccecpsspprcplgvsilidgceccmcaqqlgdncteaai 70  
OY 61 CDPHRLGYCDYSGDRPRAIGVCAQVGVGLDGVRYNNGSFQPNCKYNTCTIDGAVG 120  
DB 71 cdphrglycdysgdrpraiigvcaqvvgvlgdvrynnngsfqpnckyntctidgavg 130  
OY 121 CTPICLARPRPLMCPHRRRYSIPGHCEQWVCEDAKRRPKRTAPRDGTGAEVGEVAM 180  
DB 131 ctpiclarprplmcprrrrysipghceqwcvedakrrpkrtaprdgtgsfdaevevaw 190  
OY 181 HRNCIAYTSPMSPCSTSGGLGSTRISNVNAQCPQESRLCNLPCVDITILIKAGKK 240  
DB 191 hrnciaytspmspcstsgglgstrisnvnaqcpqesrlnclpcvditilikaqkk 250  
OY 241 CLAVYQPEASNMFTLAGCISTRSYQPKYCGVCMNRCCIPIKSKTIDVSPQDGLGFSR 300  
DB 251 clavyqp easnmftlagcistrsyqpkycgvcmnrccipksktidvspqdgldgfsr 310  
OY 301 OVLWLNACFCNLSCRNPDIFADLESYPDFSEIAN 335  
DB 311 qvlwlnacfcnlscrnpdifadlesypdfseian 345

RESULT 7

ID Y17645 standard; Protein; 367 AA.

XX Y17645;

DT 06-AUG-1999 (first entry)

DE Human WISP-1 variant protein SEQ ID NO:8.

XX WNT-1 induced secreted protein; WISP-1; WISP-2; WISP-3; CTGF; tumour;  
KW connective tissue growth factor; cancer; melanoma; arteriosclerosis;  
KW leukaemia; lymphoid malignancy; haematopoiesis-related disorder;  
KW tissue-growth disorder; skin disorder; desmoplasia; fibrotic lesion;  
KW kidney disorder; bone-related disorder; osteoporosis; trauma; burn;  
KW connective tissue disorder; catabolic state; inflammation;  
KW testicular-related disorder; angiogenesis; immunological disorder.

XX Synthetic.

OS Homo sapiens.

PN MO9921998-A1.

PD 06-MAY-1999.

PF 29-OCT-1998; 98WO-US22991.

PR 14-APR-1998; 98US-0081695.

PR 29-OCT-1997; 97US-0063704.

PR 03-FEB-1998; 98US-0073612.

XX (GETH ) GENENTECH INC.

PI Botstein DA, Cohen RL, Goddard A, Gurney AL, Hillan K;

PI Lawrence DA, Levine AJ, Penhica D, Roy MA, Wood WI;

XX WPI; 1999-337420/28.

XX New isolated Wnt-1 induced secreted polypeptides, WISP-1, 2 and 3  
PT Claim 6; Page 168-169; 284pp; English.

XX The present invention describes Wnt-1 induced secreted polypeptides,  
CC WISP-1, 2 and 3. The novel WISP polypeptides, designated WISP-1, WISP-2  
CC and WISP-3 have homology to connective tissue growth factor (CTGF).  
CC Products from the present invention can be used to treat WISP-related  
CC disorders such as breast, ovarian, and colon cancer or melanoma. The  
CC products can be used to treat arteriosclerosis. The products can also be  
CC used to treat other diseases e.g. benign and malignant tumours,  
CC leukaemia and lymphoid malignancies, neuronal, glial, astrocytal,  
CC hypothalamic and other glandular, macrophagal, epithelial, stromal, and  
CC blastocoealic disorders, haematopoiesis-related disorders, tissue-growth  
CC disorders, skin disorders, desmoplasia, fibrotic lesions, kidney  
CC disorders, bone-related disorders such as osteoporosis, trauma such as  
CC burns, incisions, and other wounds, connective tissue disorders,  
CC catabolic states, testicular-related disorders, and inflammatory,  
CC angiogenic and immunologic disorders including arteriosclerosis. The  
CC products can also be used for detection and diagnosis especially of  
CC individuals with neoplastic cell growth or proliferation. The products  
CC can be used in the production of transgenic or knock-out animals.  
CC Antibodies can be used to induce death in WISP-1, 2 or 3 overexpressing  
CC cells.

XX Sequence 367 AA;

Query Match 99.8%; Score 1934; DB 20; Length 367;  
Best Local Similarity 99.7%; Pred. No. 7.4e-138;  
Matches 334; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 DFTPAPLEDTSSRPQFCWMPCECPSPRCPLGVSILITDGCBCCKMAQQLGDNCTEAI 60  
DB 33 dftppledtssrpfqfckwpccecpsspprcplgvsilidgceccmcaqqlgdncteaai 92  
OY 61 CDPHRLGYCDYSGDRPRAIGVCAQVGVGLDGVRYNNGSFQPNCKYNTCTIDGAVG 120  
DB 93 cdphrglycdysgdrpraiigvcaqvvgvlgdvrynnngsfqpnckyntctidgavg 152  
OY 121 CTPICLARPRPLMCPHRRRYSIPGHCEQWVCEDAKRRPKRTAPRDGTGAEVGEVAM 180  
DB 153 ctpiclarprplmcprrrrysipghceqwcvedakrrpkrtaprdgtgsfdaevevaw 212  
OY 181 HRNCIAYTSPMSPCSTSGGLGSTRISNVNAQCPQESRLCNLPCVDITILIKAGKK 240  
DB 213 hrnciaytspmspcstsgglgstrisnvnaqcpqesrlnclpcvditilikaqkk 272  
OY 241 CLAVYQPEASNMFTLAGCISTRSYQPKYCGVCMNRCCIPIKSKTIDVSPQDGLGFSR 300  
DB 273 clavyqp easnmftlagcistrsyqpkycgvcmnrccipksktidvspqdgldgfsr 332  
OY 301 OVLWLNACFCNLSCRNPDIFADLESYPDFSEIAN 335  
DB 333 qvlwlnacfcnlscrnpdifadlesypdfseian 367

RESULT 8

ID Y17652 standard; Protein; 345 AA.

XX Y17652;

DT 06-AUG-1999 (first entry)

DE Human WISP-1 variant protein SEQ ID NO:21.

XX WNT-1 induced secreted protein; WISP-1; WISP-2; WISP-3; CTGF; tumour;  
KW connective tissue growth factor; cancer; melanoma; arteriosclerosis;  
KW leukaemia; lymphoid malignancy; haematopoiesis-related disorder;  
KW tissue-growth disorder; skin disorder; desmoplasia; fibrotic lesion;  
KW kidney disorder; bone-related disorder; osteoporosis; trauma; burn;

KW connective tissue disorder; catabolic state; inflammation;  
 KW testicular-related disorder; angiogenesis; immunological disorder.  
 XX Synthetic.  
 OS Homo sapiens.  
 XX  
 PN WO921998-A1.  
 XX  
 PD 06-MAY-1999.  
 XX  
 PF 29-OCT-1998; 98WO-US22991.  
 XX  
 PR 14-APR-1998; 98US-0081695.  
 XX 29-OCT-1997; 97US-0063704.  
 PR 03-FEB-1998; 98US-0073612.  
 XX  
 PA (GETH ) GENENTECH INC.  
 XX  
 PI Botstein DA, Cohen RL, Goddard A, Gurney AL, Hillan K;  
 PI Lawrence DA, Levine AJ, Pennica D, Roy MA, Wood WI;  
 XX  
 DR WPI: 1999-337420/28.  
 XX  
 PT New isolated Wnt-1 induced secreted polypeptides, WISP-1, 2 and 3  
 PS Claim 7; Page 182-183; 284pp; English.  
 XX  
 CC The present invention describes Wnt-1 induced secreted polypeptides,  
 CC WISP-1, 2 and 3. The novel WISP polypeptides, designated WISP-1, WISP-2  
 CC and WISP-3 have homology to connective tissue growth factor (CTGF).  
 CC Products from the present invention can be used to treat WISP-related  
 CC disorders such as breast, ovarian, and colon cancer or melanoma. The  
 CC products can be used to treat arteriosclerosis. The products can also be  
 CC used to treat other diseases e.g. benign and malignant tumours,  
 CC leukaemia and lymphoid malignancies, neuronal, glial, astrocytal,  
 CC hypoblastic and other glandular, macrophagal, epithelial, stromal, and  
 CC blastocoele disorders, haematopoiesis-related disorders, tissue-growth  
 CC disorders, skin disorders, desmoplasia, fibrotic lesions, kidney  
 CC disorders, bone-related disorders such as osteoporosis, trauma such as  
 CC burns, incisions, and other wounds, connective tissue disorders,  
 CC catabolic states, testicular-related disorders, and inflammatory,  
 CC angiogenic and immunologic disorders including arteriosclerosis. The  
 CC products can also be used for detection and diagnosis especially of  
 CC individuals with neoplastic cell growth or proliferation. The products  
 CC can be used in the production of transgenic or knock-out animals.  
 CC Antibodies can be used to induce death in WISP-1, 2 or 3 overexpressing  
 CC cells.  
 XX  
 SQ Sequence 345 AA;

Query Match 99.8%; Score 1933; DB 20; Length 345;  
 Best Local Similarity 99.4%; Pred. No. 8.3e-138;  
 Matches 333; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 DFTFPALEDTSSRPQFCWPCCEPPSPRCPGLVSLITDGECCCKMAQQLGDNCTEAAI 60  
 DB 11 dftfpaledtssrpfqfckwpcceppsprrcpplgysliltdegeccckmaeqqlgdncteaai 70  
 QY 61 CDPRRGIXCYDSGRPRRYAYGCAVYGVGVLDGVRYNKNGSQPMCKRYKCTCIDAVG 120  
 DB 71 cdprrgixcydsgrprryaygcvaygvvcvldgvrnyngsqfpmckrykctcidavg 130  
 QY 121 CTPLCLVPRPRLMCPHRRYSIRGHCCEDQWVEDAKRPRKTPARPTGAFDAVEYEAW 180  
 DB 131 ctplclvprprlmcphrrysirghccedqwcvedakrprktparptdgafdaveyeaw 190  
 QY 181 HRCNIATSPWSPGSCGAGVSTRISVNAQCWPEBSRLCNLRPCVDIHTLIIKAGKK 240  
 DB 191 hrcniatspwspgscgagvstrisvnaqcwpebsrlcnlrpcvdihtlikagkk 250  
 QY 241 CLAVYQPEASNMFTLACISIRSYQPKYCGVCMNRCIPKSKTIVSFQCPGLGFSR 300  
 DB 311 clavyqpeasnmftlacisirsyqpkycgvcmnrcipksktivsftqcpdglgfsr 310

DB 251 clavyqpeasnmftlacisirsyqpkycgvcmnrcipksktivsftqcpdglgfsr 310  
 QY 301 QVLMINACFCNLSCRNPNDIFADLESTPDRSEIN 335  
 DB 311 qvlinacfcnlscrnpndifadlesypdrselan 345

RESULT 9  
 Y17653  
 ID Y17653 standard; Protein; 367 AA.  
 XX  
 AC Y17653;  
 XX  
 DT 06-AUG-1999 (first entry)  
 XX  
 DE Human WISP-1 variant protein SEQ ID NO:22.  
 XX  
 KW WNT-1 induced secreted protein; WISP-1; WISP-2; WISP-3; CTGF; tumour;  
 KW connective tissue growth factor; cancer; melanoma; arteriosclerosis;  
 KW leukaemia; lymphoid malignancy; haematopoiesis-related disorder;  
 KW tissue-growth disorder; skin disorder; desmoplasia; fibrotic lesion;  
 KW kidney disorder; bone-related disorder; osteoporosis; trauma; burn;  
 KW connective tissue disorder; catabolic state; inflammation;  
 KW testicular-related disorder; angiogenesis; immunological disorder.  
 XX  
 OS Synthetic.  
 OS Homo sapiens.  
 XX  
 PN WO921998-A1.  
 XX  
 PD 06-MAY-1999.  
 XX  
 PF 29-OCT-1998; 98WO-US22991.  
 XX  
 PR 14-APR-1998; 98US-0081695.  
 XX 29-OCT-1997; 97US-0063704.  
 PR 03-FEB-1998; 98US-0073612.  
 XX  
 PA (GETH ) GENENTECH INC.  
 XX  
 PI Botstein DA, Cohen RL, Goddard A, Gurney AL, Hillan K;  
 PI Lawrence DA, Levine AJ, Pennica D, Roy MA, Wood WI;  
 XX  
 DR WPI: 1999-337420/28.  
 XX  
 PT New isolated Wnt-1 induced secreted polypeptides, WISP-1, 2 and 3  
 PS Claim 7; Page 183-184; 284pp; English.  
 XX  
 CC The present invention describes Wnt-1 induced secreted polypeptides,  
 CC WISP-1, 2 and 3. The novel WISP polypeptides, designated WISP-1, WISP-2  
 CC and WISP-3 have homology to connective tissue growth factor (CTGF).  
 CC Products from the present invention can be used to treat WISP-related  
 CC disorders such as breast, ovarian, and colon cancer or melanoma. The  
 CC products can be used to treat arteriosclerosis. The products can also be  
 CC used to treat other diseases e.g. benign and malignant tumours,  
 CC leukaemia and lymphoid malignancies, neuronal, glial, astrocytal,  
 CC hypoblastic and other glandular, macrophagal, epithelial, stromal, and  
 CC blastocoele disorders, haematopoiesis-related disorders, tissue-growth  
 CC disorders, skin disorders, desmoplasia, fibrotic lesions, kidney  
 CC disorders, bone-related disorders such as osteoporosis, trauma such as  
 CC burns, incisions, and other wounds, connective tissue disorders,  
 CC catabolic states, testicular-related disorders, and inflammatory,  
 CC angiogenic and immunologic disorders including arteriosclerosis. The  
 CC products can also be used for detection and diagnosis especially of  
 CC individuals with neoplastic cell growth or proliferation. The products  
 CC can be used in the production of transgenic or knock-out animals.  
 CC Antibodies can be used to induce death in WISP-1, 2 or 3 overexpressing  
 CC cells.  
 XX  
 SQ Sequence 367 AA;



Query Match 99.8%; Score 1933; DB 20; Length 367;  
 Best Local Similarity 99.4%; Pred. No. 8,8e-138;  
 Matches 333; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

1 DFTPALEDTSSRPQFCCKWPCBPSPRCPLGVSILITDGCCECKKCAQQLGDNCTEAAI 60  
 33 dftppledtssrpfckwpcbpsprrcplgvsilidgcceckkcaqqlgdncteeai 92  
 Db 33 dftppledtssrpfckwpcbpsprrcplgvsilidgcceckkcaqqlgdncteeai 92

61 CDPRHGLVCDYSGDRPRAIGVCAOVGVGLDGVRRVNGSFOFPNCKYKNCITIGAVG 120  
 93 cdprhglvcdysgdrpraigvcaovgvglldgvrrvngsfofpnckynctidgavg 152  
 Db 93 cdprhglvcdysgdrpraigvcaovgvglldgvrrvngsfofpnckynctidgavg 152

121 CPRLCLRVRRPRLMCPHRRVRSIPGHCCBQWCEDDAKRPRTARPDGAFNAGEVEM 180  
 133 cprlclrvrrprrlmcprrvrsipghccbwceddakrrprtardgafnagevem 212  
 Db 133 cprlclrvrrprrlmcprrvrsipghccbwceddakrrprtardgafnagevem 212

181 HRNCIAYTSPMSPCSTSCGLGVSTRISNVNAQCMPEQESRLCNLRPCVDVHTLLKAGK 240  
 213 hrnciaytspmspcstscglgvstrisnvnaqcmpegesrlcnlrvcdvhtllkagk 272  
 Db 213 hrnciaytspmspcstscglgvstrisnvnaqcmpegesrlcnlrvcdvhtllkagk 272

241 CLAVYQPEASMNFTLAGCISTRSYQPKYCGVCMNRCIPYKSKTIDVSFQCPDGLGFSR 300  
 273 clavyqpeasnmftlagcistrsyqpkycgvcmnrcipkysktidvsfqpdlgfsr 332  
 Db 273 clavyqpeasnmftlagcistrsyqpkycgvcmnrcipkysktidvsfqpdlgfsr 332

301 OVLWINACFCNLSCRNPNDIFADLESYPDSRIAN 335  
 333 qvlwinacfcnlscrnndifadlesypdsriean 367  
 Db 333 qvlwinacfcnlscrnndifadlesypdsriean 367

RESULT 10  
 Y17646  
 ID Y17646 standard; Protein: 345 AA.  
 AC Y17646;  
 DT 06-AUG-1999 (first entry)  
 DE Mouse putative mature WISP-1 protein SEQ ID NO:11.

WNT-1 induced secreted protein; WISP-1; WISP-2; WISP-3; CTGF; tumour;  
 connective tissue growth factor; cancer; melanoma; arteriosclerosis;  
 leukaemia; lymphoid malignancy; haematopoiesis-related disorder;  
 tissue-growth disorder; skin disorder; desmoplasia; fibrotic lesion;  
 kidney disorder; bone-related disorder; osteoporosis; trauma; burn;  
 connective tissue disorder; catabolic state; inflammation;  
 testicular-related disorder; angiogenesis; immunological disorder.

Mus sp.  
 W09921998-A1.  
 PD 06-MAY-1999.  
 PF 29-OCT-1998; 98WO-US22991.  
 PR 14-APR-1998; 98US-0081695.  
 PR 29-OCT-1997; 97US-0063704.  
 PR 03-FEB-1998; 98US-0073612.  
 PA (GETH ) GENENTECH INC.  
 PI Botstein DA, Cohen RL, Goddard A, Gurney AL, Hillan K;  
 PI Lawrence DA, Levine AJ, Pennica D, Roy MA, Wood WI;  
 DR WPI; 1999-337420/28.  
 PT New isolated wnt-1 induced secreted polypeptides, WISP-1, 2 and 3  
 PS Claim 9; Page 172-173; 284pp; English.  
 CC The present invention describes wnt-1 induced secreted polypeptides,  
 CC WISP-1, 2 and 3. The novel WISP polypeptides, designated WISP-1,  
 CC and WISP-3 have homology to connective tissue growth factor (CTGF).

CC Products from the present invention can be used to treat WISP-related  
 CC disorders such as breast, ovarian, and colon cancer or melanoma. The  
 CC products can be used to treat arteriosclerosis. The products can also be  
 CC used to treat other diseases e.g. benign and malignant tumours,  
 CC leukaemia and lymphoid malignancies, neuronal, glial, astrocytic,  
 CC hypothalamic and other glandular, macrophagel, epithelial, stromal, and  
 CC blastocoeleic disorders, haematopoiesis-related disorders, tissue-growth  
 CC disorders, skin disorders, desmoplasia, fibrotic lesions, kidney  
 CC disorders, bone-related disorders such as osteoporosis, trauma such as  
 CC burns, incisions, and other wounds, connective tissue disorders,  
 CC catabolic states, testicular-related disorders, and inflammatory,  
 CC angiogenic and immunologic disorders including arteriosclerosis. The  
 CC products can also be used for detection and diagnosis especially of  
 CC individuals with neoplastic cell growth or proliferation. The products  
 CC can be used in the production of transgenic or knock-out animals.  
 CC Antibodies can be used to induce death in WISP-1, 2 or 3 overexpressing  
 CC cells.

Sequence 345 AA:  
 SQ

Query Match 86.5%; Score 1675; DB 20; Length 345;  
 Best Local Similarity 85.3%; Pred. No. 1,8e-118;  
 Matches 285; Conservative 20; Mismatches 29; Indels 0; Gaps 0;

2 FTRPALEDTSSRPQFCCKWPCBPSPRCPLGVSILITDGCCECKKCAQQLGDNCTEAAI 61  
 12 ftrpaledtssrpfckwpcbpsprrcplgvsilidgcceckkcaqqlgdncteeai 61  
 Db 12 ftrpaledtssrpfckwpcbpsprrcplgvsilidgcceckkcaqqlgdncteeai 61

62 DPHRGLVCDYSGDRPRAIGVCAOVGVGLDGVRRVNGSFOFPNCKYKNCITIGAVG 121  
 72 dphrhlvcdysgdrpraigvcaovgvglldgvrrvngsfofpnckynctidgavg 131  
 Db 72 dphrhlvcdysgdrpraigvcaovgvglldgvrrvngsfofpnckynctidgavg 131

122 TPCLCLRVRRPRLMCPHRRVRSIPGHCCBQWCEDDAKRPRTARPDGAFNAGEVEM 181  
 132 tpclclrvrrprrlmcprrvrsipghccbwceddakrrprtardgafnagevem 191  
 Db 132 tpclclrvrrprrlmcprrvrsipghccbwceddakrrprtardgafnagevem 191

182 HRNCIAYTSPMSPCSTSCGLGVSTRISNVNAQCMPEQESRLCNLRPCVDVHTLLKAGK 241  
 192 hrnciaytspmspcstscglgvstrisnvnaqcmpegesrlcnlrvcdvhtllkagk 251  
 Db 192 hrnciaytspmspcstscglgvstrisnvnaqcmpegesrlcnlrvcdvhtllkagk 251

242 LAVYQPEASMNFTLAGCISTRSYQPKYCGVCMNRCIPYKSKTIDVSFQCPDGLGFSR 301  
 252 lavyqpeasnmftlagcistrsyqpkycgvcmnrcipkysktidvsfqpdlgfsr 311  
 Db 252 lavyqpeasnmftlagcistrsyqpkycgvcmnrcipkysktidvsfqpdlgfsr 311

302 VLMINACFCNLSCRNPNDIFADLESYPDSRIAN 335  
 312 vlmnacfcnlscrnndifadlesypdsriean 345  
 Db 312 vlmnacfcnlscrnndifadlesypdsriean 345

RESULT 11  
 Y17647  
 ID Y17647 standard; Protein: 367 AA.  
 AC Y17647;  
 DT 06-AUG-1999 (first entry)  
 DE Mouse WISP-1 protein SEQ ID NO:12.

WNT-1 induced secreted protein; WISP-1; WISP-2; WISP-3; CTGF; tumour;  
 connective tissue growth factor; cancer; melanoma; arteriosclerosis;  
 leukaemia; lymphoid malignancy; haematopoiesis-related disorder;  
 tissue-growth disorder; skin disorder; desmoplasia; fibrotic lesion;  
 kidney disorder; bone-related disorder; osteoporosis; trauma; burn;  
 connective tissue disorder; catabolic state; inflammation;  
 testicular-related disorder; angiogenesis; immunological disorder.

Mus sp.  
 W09921998-A1.  
 PD 06-MAY-1999.



XX 29-OCT-1998; 98WO-US22991.  
 PF 14-APR-1998; 98US-0081695.  
 PR 29-OCT-1997; 97US-0063704.  
 PR 03-FEB-1998; 98US-0073612.

XX (GETH ) GENENTECH INC.

PI Botstein DA, Cohen RL, Goddard A, Gurney AL, Hillan K;  
 PI Lawrence DA, Levine AJ, Pennica D, Roy MA, Wood WI;

DR WPI: 1999-337420/28.  
 DR N-PSDB; X76484.

PT New isolated Wnt-1 induced secreted polypeptides, WISP-1, 2 and 3  
 PS Claim 9; Page 173-174; 284pp; English.

XX The present invention describes Wnt-1 induced secreted polypeptides,  
 CC WISP-1, 2 and 3. The novel WISP polypeptides, designated WISP-1, WISP-2  
 CC and WISP-3 have homology to connective tissue growth factor (CTGF).  
 CC Products from the present invention can be used to treat WISP-related  
 CC disorders such as breast, ovarian, and colon cancer or melanoma. The  
 CC products can be used to treat arteriosclerosis. The products can also be  
 CC used to treat other diseases e.g. benign and malignant tumors,  
 CC leukaemia and lymphoid malignancies, neuronal, glial, astrocytal,  
 CC hypothalamic and other glandular, macrophagal, epithelial, stromal, and  
 CC blastocoeleic disorders, haematopolesis-related disorders, tissue-growth  
 CC disorders, skin disorders, desmoplasia, fibrotic lesions, kidney  
 CC disorders, bone-related disorders such as osteoporosis, trauma such as  
 CC burns, incisions, and other wounds, connective tissue disorders,  
 CC catabolic states, testicular-related disorders, and inflammatory,  
 CC angiogenic and immunologic disorders including arteriosclerosis. The  
 CC products can also be used for detection and diagnosis especially of  
 CC individuals with neoplastic cell growth or proliferation. The products  
 CC can be used in the production of transgenic or knock-out animals.  
 CC Antibodies can be used to induce death in WISP-1, 2 or 3 overexpressing  
 CC cells.

SQ Sequence 367 AA;

Query Match 86.5%; Score 1675; DB 20; Length 367;  
 Best Local Similarity 85.3%; Pred. No. 1.9e-118;  
 Matches 285; Conservative 20; Mismatches 29; Indels 0; Gaps 0;

QY 2 FTAPLEDTSSRPOFCMKPECEPPSPRCLVGLITDGCCECKMCAQOLGDNCTEAATC 61  
 DB 34 fcpapleettitpctckpcepqsprrcpilvslldgceccckicqqlgdncteaatc 93  
 QY 62 DPHRGLYCDVSGDRPRYAVCAQVGVGCVLDGVRRYNNQSGFOPNCKYCTCIDGAVGC 121  
 DB 94 dphrglycdvsgdrrpryavcaqvvgvlgvrryngsfopnckyctcidgavgc 153  
 QY 122 TPLCLKVRPRRLMCPHRRKRSIRGHCCEQWVCEDDAKRPKRTAPRTGAFDAVGEVAMH 181  
 DB 154 tplclskvrprrlmcprrrksirgcccqvwvceddakrrpkrtaprtgafdaevamh 213  
 QY 182 RMCIAVTSPPSPGSGTSGVTRISNVNQAQWPEQESRLCNIRPCVDVHTLTKAKKCC 241  
 DB 214 rmciavtspspgstsgtsgvtrisnvnqaqwpeeqesrlcnirpcvdvhtltkakkc 273  
 QY 242 LAVYQPEASNNFTLAGICSTRSYQPKYGVGMNRRCCIPYKSKTIDVFSQCDGLGFSRQ 301  
 DB 274 lavyqpeasnnftlagicstrsyqpkycvgtmnnrrccipykstktdvfsqcdglgfsrq 333  
 QY 302 VLVINMCFCLSCGRNPRDIFADLESYRDESEIAN 335  
 DB 334 vlvinnmcfclscgrnprdifadlesyrdeeseian 367

RESULT 12

R79964  
 ID R79964 standard; Protein; 349 AA.

XX AC R79964;

DT 12-JUN-1996 (first entry)

XX DE Connective tissue growth factor.

XX KW Connective tissue growth factor; CTGF; wound healing; vulnery;  
 KW cell proliferation; cancer; fibrosis; atherosclerosis; diagnosis  
 KW therapy; mitogen.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

FT Modified-site 28 /label= N-glycosylation\_site  
 FT Modified-site 225 /label= N-glycosylation\_site

XX US5408040-A.

XX 18-APR-1995.

XX 30-AUG-1991; 91US-0752427.

XX 30-AUG-1991; 91US-0752427.

XX 14-DEC-1993; 93US-0167628.

XX (USF-) UNIV SOUTH FLORIDA.

XX Bradham DM, Grotendorst GR;

XX WPI: 1995-161147/21.

XX N-PSDB; T04226.

PT New connective tissue growth factor - used to develop prod. for  
 PT wound healing and for diagnosis and therapy of cell proliferative  
 PT disorders.

PS Claim 1; Column 19-20; 12pp; English.

XX Novel human connective tissue growth factor (CTGF) (R79964)  
 CC is related immunologically and biologically to platelet-derived  
 CC growth factor (PDGF), but is the product of a distinct gene.  
 CC CTGF is mitogenic and also a chemotactic agent for cells. It is  
 CC produced by endothelial and fibroblastic cells, and probably acts  
 CC as a growth factor in wound healing. Recombinant CTGF can be obtd.  
 CC by expression of cDNA clone DB6OR32 (T04226) in transformed host  
 CC cells. It is used to accelerate wound healing, and to raise  
 CC antibodies useful in detecting disorders associated with overgrowth  
 CC of cells, such as cancer, fibrotic diseases and atherosclerosis.

SQ Sequence 349 AA;

Query Match 44.9%; Score 869.5; DB 16; Length 349;  
 Best Local Similarity 45.6%; Pred. No. 4.6e-58;  
 Matches 156; Conservative 53; Mismatches 98; Indels 35; Gaps 8;

QY 12 SRP---QFCMKPECEPPSP-PRCPGLVSLITDGCCECKMCAQOLGDNCTEAATDPRGL 67  
 DB 21 srpvgqncsgpcrcpdepaprcpavslvldgsgccrrcvaqqlgclcterdpdpbkyl 80  
 QY 68 YCDVSGDRPRYAVCAQVGVGCVLDGVRRYNNQSGFOPNCKYCTCIDGAVGCTPLC-L 126  
 DB 81 ycdvsgdrpryavcaqvvgvlgvrryngsfopnckyctcidgavgcplcsm 139  
 QY 127 RVPRRLMCPHRRKRSIRGHCCEQWVCEDDAKRPKRTAPRTGAFDAVGEVAMH---- 182  
 DB 140 dvrlpspdcprfprtrvklpjkccceewcde-----pkdq---tvvgalaaayrltcd 187

QY 183 -----NCIATSPWSPCSTSCGLGVSTRISNVNNAOCPEDESRCLNRPDDVDIH 232  
 Db fgpdpmtiranclygtltewscsktcgmgjstrvtnndascrilekqsrilcmvrpceadle 247  
 OY 233 TLIRAGKRCIAVYOPEASGMNFTLACICSTRSYOPKVCVCMNDRCICPYKSKTIDVSPQC 292  
 Db 248 enlkkgkckltpklskplkelsgtctsmktyrakfcgvcldgcrctcphtltpvckfc 307  
 QY 293 PDGLGFSRQVIMINACFCNLSCRNPNDFADL---ESYPDFS 331  
 Db 308 pdgwmkmmfiktckachyncpgndlfeslyrkmygdma 349  
 RESULT 13  
 W09089  
 ID W09089 standard; Protein: 349 AA.  
 AC W09089;  
 XX 26-APR-1997 (first entry)  
 DE Human connective tissue growth factor.  
 KW Connective tissue growth factor; CTGF; mitogen; cell proliferation;  
 KW wound healing; cancer; tumour; fibrosis; glaucoma; atherosclerosis;  
 KW scleroderma; arthritis; cirrhosis; scar; diagnosis; therapy.  
 OS Homo sapiens.  
 FH Key location/qualifiers  
 FT Modified-site 28 /label= Glycosylation  
 FT /note= "potential N-glycosylation site"  
 FT Modified-site 225 /label= Glycosylation  
 FT /note= "potential N-glycosylation site"  
 FN W09638172-A1.  
 XX 05-DEC-1996.  
 PD 31-MAY-1996; 96WO-US08140.  
 XX 31-MAY-1996; 96WO-US08140.  
 PR (UYSF-) UNIV SOUTH FLORIDA.  
 PA Bradham DM, Grotendorst GR;  
 PI WPI: 1997-042659/04.  
 DR N-PSDB; T45360;  
 DR N-PSDB; T58534.  
 XX Connective tissue growth factor coding sequence and protein - used  
 PT in the treatment of proliferative disorders and to accelerate wound  
 PT healing  
 PS Claim 19; Page 50-52; 76pp; English.  
 XX Novel human connective tissue growth factor (CTGF) (W09089) is a  
 CC PDGF-immunorelated protein that may play a significant role in the  
 CC normal development, growth and repair of human tissue and probably  
 CC functions as a growth factor in wound healing. CTGF may be involved  
 CC in diseases in which there is an overgrowth of connective tissue  
 CC cells, such as cancer, tumour formation and growth, fibrotic  
 CC diseases (e.g. pulmonary fibrosis, kidney fibrosis, glaucoma) and  
 CC atherosclerosis. Recombinant CTGF can be produced in transformed  
 CC host cells utilising a cDNA clone isolated from a HUVEC library.  
 CC It can be used to accelerate wound healing. CTGF inhibitors can be  
 CC used to treat atherosclerosis and fibrotic diseases such as  
 CC scleroderma, arthritis, liver cirrhosis, and scarring.  
 XX Sequence 349 AA;

Query Match 44.9%; Score 869.5; DB 18; Length 349;  
 Best Local Similarity 45.6%; Pred. No. 4.6e-58;  
 Matches 156; Conservative 53; Mismatches 98; Indels 35; Gaps 8;  
 OY 12 SRP---QFCWKPCBEPSPD-PRCPGLVSLITDGCRCCKMACAOLODNCTEAATCPHRL 67  
 Db 21 srpavgnscspcrpcdpdpaprcpavvalvldgcgcrcvcaqlgclctcrpdpbhkyl 80  
 OY 68 YCDYSGDRPRRATIGCAQVGVCLDGVRRYNNNGSPQBNCKYNTCTIDGAVGCPRLC-L 126  
 Db 81 fcdt-gspanrklyvctckdgcpcifgltvryrsesfgsclyqctcldgavagmplcam 139  
 OY 127 RYRPPRLMCPHPRRVSIDGHCCEQWVCEDDAKRPRTAPRDTGAPDAVEEAMHR--- 182  
 Db 140 dvrlpspcpfrvrkklpgkceewvde-----pkdq---tlvgsalaeayrlcdt 187  
 OY 183 -----NCIATSPWSPCSTSCGLGVSTRISNVNNAOCPEDESRCLNRPDDVDIH 232  
 Db 188 fgpdpmtiranclygtltewscsktcgmgjstrvtnndascrilekqsrilcmvrpceadle 247  
 OY 233 TLIRAGKRCIAVYOPEASGMNFTLACICSTRSYOPKVCVCMNDRCICPYKSKTIDVSPQC 292  
 Db 248 enlkkgkckltpklskplkelsgtctsmktyrakfcgvcldgcrctcphtltpvckfc 307  
 OY 293 PDGLGFSRQVIMINACFCNLSCRNPNDFADL---ESYPDFS 331  
 Db 308 pdgwmkmmfiktckachyncpgndlfeslyrkmygdma 349  
 RESULT 14  
 ID W11302  
 W11302 standard; Protein: 349 AA.  
 AC W11302;  
 XX 18-MAR-1997 (first entry)  
 DE Connective tissue growth factor.  
 KW Connective tissue growth factor; CTGF; human; connective tissue cell;  
 KW proliferative disease; platelet-derived growth factor; PDGF; development;  
 KW tissue growth; repair; umbilical vein endothelial cell; HUVE cell;  
 KW antibody; wound healing; cancer; fibrotic disease; atherosclerosis;  
 KW inhibitor; protease degradation; growth factor; therapy.  
 OS Homo sapiens.  
 FH Homo sapiens.  
 FN US5585270-A.  
 DR US5585270-A.  
 DR 17-DEC-1996.  
 PD 30-AUG-1991; 91US-0752427.  
 XX 30-AUG-1991; 91US-0752427.  
 PR 30-AUG-1991; 91US-0752427.  
 PR 14-DEC-1993; 93US-0167628.  
 PR 10-FEB-1995; 95US-0386680.  
 XX (UYSF-) UNIV SOUTH FLORIDA.  
 PA Bradham DM, Grotendorst GR;  
 PI WPI: 1997-051180/05.  
 DR N-PSDB; T51234.  
 XX New nucleic acid encoding connective tissue growth factor - useful  
 PT for accelerating wound healing, also for diagnosis and treatment of  
 PT proliferative disease  
 PS Claim 9; Column 15-18; 11pp; English.  
 XX This sequence represents the human connective tissue growth factor

CC (CTGF). CTGF is related immunologically and biologically to  
 CC platelet-derived growth factor (PDGF), but is encoded by an unrelated  
 CC gene. CTGF is thought to play a significant role in the normal  
 CC development, growth, and repair of human tissue, similarly to PDGF. The  
 CC cDNA encoding this sequence was isolated by screening a cDNA library from  
 CC human umbilical vein endothelial (HUVE) cells with anti-PDGF antibodies.  
 CC CTGF can be used to accelerate wound healing. Also, elevated levels of  
 CC CTGF may be diagnostic of proliferative diseases involving outgrowth of  
 CC connective tissue cells, such as cancer, fibrotic disease and  
 CC atherosclerosis. All of these diseases can be treated with reagents  
 CC reactive with CTGF, such as antibodies (which can also serve as assay  
 CC reagents). Antisense nucleic acids, and ribozymes could also be used to  
 CC inhibit CTGF production. The advantage with using CTGF is that it is  
 CC more stable, and less susceptible to protease degradation than PDGF, and  
 CC other growth factors involved in wound healing. This is believed to be  
 CC due to the high Cys content.

CC Sequence 349 AA:

Query Match 44.9%; Score 869.5; DB 18; Length 349;  
 Best Local Similarity 45.6%; Pred. No. 4.6e-58;  
 Matches 156; Conservative 53; Mismatches 98; Indels 35; Gaps 8;

QY 12 SRP---QFCMKPCPCPPSP-PRCPGLVSLITDGECCCKMCAQQLGDMCTEALICDPHRL 67  
 DB 21 srpavqncsgpcrcpdepaprcpavslvldgqccrvcaakqglgclcterdpdpkhgyl 80  
 QY 68 YCDYSGDRPRKATGVCAQVAVGCVLDGVRNNQOSFQPNCKYKCTCIDGAVGCTPLC-L 126  
 DB 81 fcdt-gepanrklyvctackdgapecifgtyrsgesfsgskycgtcldgavgcmlcsm 139  
 QY 127 RVRRPRLMCPHRRVSIPIGHCCEDQWCEDDAKRRKTAAPDGTGAFDAVGEVAMHR--- 182  
 DB 140 dvrlpspcpfprrvklpgkceewcde-----pkdq---tvvgpalaayrldet 187  
 QY 183 -----NCIAYTSPWSPSTSCGLGVSTRISNNVAQCPQESRLNLRPCVDI 232  
 DB 188 fgpdpmlrancelvgtewasacktcgmglatrvtnnascrlekqslcmvrpcedle 247  
 QY 233 TLIRAGKCLAVYQPEASMNFTLAGCISTRSYQRYKCYGVCMNDNRCIPIYKSKTIDVSFQC 292  
 DB 248 enlkkgkclrtprkiskipikfclsgctsmktyrakfcgvcldgrccphrttlipvefk 307  
 QY 293 PDGLGFSROYLWINACFCNLSCRNPNDIFADL---ESYPDFS 331  
 DB 308 pdgevmkmmfiktcahyncpgndlfeslyyrkmygdma 349

RESULT 15  
 ID W62084 standard; Protein; 349 AA.

AC W62084;

DT 15-SEP-1998 (first entry)

DE Human connective tissue growth factor.

KM Human: connective tissue growth factor; CTGF; PDGF; diagnosis; cancer;

KW platelet derived growth factor; ameliorating cell proliferative disorder;

KW atherosclerosis; fibrotic disease.

OS Homo sapiens.

PN US5783187-A.

PD 21-JUL-1998.

PF 11-SEP-1996; 96US-0712302.

PR 30-AUG-1991; 91US-0752427.

PR 14-DEC-1993; 93US-0167628.

PR 11-SEP-1996; 96US-0712302.

PA (UYSF-) UNIV SOUTH FLORIDA.

PI Bradham DM, Grotendorst GR;

DR WPI: 1998-426958/36.

DR N-PSDB; V38085.

PT Ameliorating cell proliferative disorder associated with connective  
 PT tissue growth factor - comprises the administration of an antibody  
 PT which binds to connective tissue growth factor and not to  
 PT platelet-derived growth factor

PS Example 6; Column 17-20; 11pp; English.

CC A method has been developed for ameliorating a cell proliferative  
 CC disorder associated with connective tissue growth factor (CTGF). The  
 CC method comprises the administration of an antibody or its fragment that  
 CC binds to CTGF and not to platelet-derived growth factor (PDGF), to the  
 CC site of the disorder. CTGF is related immunologically and biologically  
 CC to PDGF. The present sequence represents CTGF. The method is used to  
 CC treat conditions involving the overgrowth of connective tissue cells  
 CC such as cancer, atherosclerosis and other fibrotic diseases.

CC Sequence 349 AA:

Query Match 44.9%; Score 869.5; DB 19; Length 349;  
 Best Local Similarity 45.6%; Pred. No. 4.6e-58;  
 Matches 156; Conservative 53; Mismatches 98; Indels 35; Gaps 8;

QY 12 SRP---QFCMKPCPCPPSP-PRCPGLVSLITDGECCCKMCAQQLGDMCTEALICDPHRL 67  
 DB 21 srpavqncsgpcrcpdepaprcpavslvldgqccrvcaakqglgclcterdpdpkhgyl 80  
 QY 68 YCDYSGDRPRKATGVCAQVAVGCVLDGVRNNQOSFQPNCKYKCTCIDGAVGCTPLC-L 126  
 DB 81 fcdt-gepanrklyvctackdgapecifgtyrsgesfsgskycgtcldgavgcmlcsm 139  
 QY 127 RVRRPRLMCPHRRVSIPIGHCCEDQWCEDDAKRRKTAAPDGTGAFDAVGEVAMHR--- 182  
 DB 140 dvrlpspcpfprrvklpgkceewcde-----pkdq---tvvgpalaayrldet 187  
 QY 183 -----NCIAYTSPWSPSTSCGLGVSTRISNNVAQCPQESRLNLRPCVDI 232  
 DB 188 fgpdpmlrancelvgtewasacktcgmglatrvtnnascrlekqslcmvrpcedle 247  
 QY 233 TLIRAGKCLAVYQPEASMNFTLAGCISTRSYQRYKCYGVCMNDNRCIPIYKSKTIDVSFQC 292  
 DB 248 enlkkgkclrtprkiskipikfclsgctsmktyrakfcgvcldgrccphrttlipvefk 307  
 QY 293 PDGLGFSROYLWINACFCNLSCRNPNDIFADL---ESYPDFS 331  
 DB 308 pdgevmkmmfiktcahyncpgndlfeslyyrkmygdma 349

Search completed: February 16, 2001, 16:13:31  
 Job time: 42 sec

---



```

RESULT      2
A40578
beta IG-M2 protein precursor - mouse
C:Species: Mus musculus (house mouse)
C:Date: 06-Mar-1992 #sequence_revision 06-Mar-1992 #ext_change 05-Nov-1999
C:Accession: A40578
R:Burner, A.; Chinn, J.; Neubauer, M.; Purchio, A.F.
DNA Cell Biol. 10, 293-300, 1991
A:Title: Identification of a gene family regulated by transforming growth factor-beta
A:Reference number: A40578; MUID:91229699
A:Accession: A40578
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-348 <BRU>
A:Cross-references: GB:M80263; NID:g201945; PIDN:AAA7135.1; PID:g201946

Query Match      43.1%; Score 835.5; DB 2; Length 348;
Best Local Similarity 44.4%; Pred. No. 7.3e-55;
Matches 152; Conservative 55; Mismatches 100; Indels 35; Gaps 8;

```

```

QY      12  SRP - FQCKAPRC - PPSPPRCGLSLITDGCCECKMAQOGLDCTEALIIDPRGL  67
Db      20  TRPATGDCSCAQCCAAEAPAHPCPAGVSLYLDGCGCCRYCAKQJGELCTEBDPCHKGL  79

QY      68  YCDISGDRPRAIATGCAVAGVGCVLDSGVRNNQSGSPMCKYTKCTSIDAVGCTPLC-L  126
Db      80  FCDP - GSPANKRIDVCTAKDAPCFVGGSVYRSESSFOCKTCTLDGAVGVLPLCSM  138

QY      127  RVPRRLMCPRRRVSIPGHCCQMWCEDDAKRRPRKTAAPRDGAFDAVGEEYEAHR- ---  182
Db      139  DVRLPSPPDCPPRRVKLPGKCCEMWCD- ----PKR- -TAVGALAAVRIEDT  186

QY      183  -----NCIATVSPSCSTSCGLGVSTRISNNNAQCMRPQESRLCLNRCQVDYDIH  232
Db      187  FGDPPTMMRACILVQTEWMSACSKTCGISTRYTNTNPTCRLEKQSRILCMVRCSEADLE  246

QY      233  TLIRAGKCLAVOYPEASMNFTLAGCISTRSYDPRYGCVCMDNRCCLPYKSTIDVAFSQC  292
Db      247  ENIKGKKCIRTPIAKRVKFPELSGCISVTKYRAKFCGVCTDRCGCRPHRTTLPVEFKC  306

QY      293  PDGLGFSQVLIINACFCNLSCRPNIDFADL- ---ESTPPES  331
Db      307  PDGELIMKNMAMFICTCAHYNCPGDNDIFSLYIRKMTGDMA  348

```

```

RESULT      3
A53228
fisp-12 protein precursor - mouse
C:Species: Mus musculus (house mouse)
C:Date: 19-May-1994 #sequence_revision 19-May-1994 #text_change 05-Nov-1999
C:Accession: A53228
R:Ryseck, R.P.; MacDonald-Bravo, H.; Mattei, M.G.; Bravo, R.
Cell Growth Differ. 2, 225-233, 1991
A:Title: Structure, mapping, and expression of fisp-12, a growth factor-inducible gene
A:Reference number: A53228; MIMD:91363290
A:Accession: A53228
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-348 <RYS>
A:Cross-references: GB:M70641; NID:g193313; PIDN:AAA37627.1; PID:g193314
C:Genetics:
C:Gene: fisp-12

```

```

Query Match      42.98;  Score 831.5;  DB 2;  Length 348;
Best Local Similarity  44.25;  Pred. No. 1,4e-54;
Matches 151;  Conservative 56;  Mismatches 100;  Indels 35;  Gaps 8;

QY      12 SRP---QPCCKPCCG-PPSPRCPLGVSLLTIDGECCKMCAQOLGDMCTEAATDPPRGL 67
      :||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |

```

```

Db      20  TRPATGGDSCAACCACCAEAALPHCPAGASLVLDGCGCCRCAPCAKOLGELCTERPDCPHKL  79
QY      68  YCDSSGSRPRATAYAGVCAQVGVGCVLDGVRVYNNNGSFPQNNKNCIDGAVGCPILC-L  126
Db      80  FCDP-GSPANRKIRISVCTAKAGCAPCVGSGSYRSGESFQSSCKTCTCLEGAGVCPILSM  138
QY     127  RVRPRLMCPHPRRVSIPGHCCQEWVCEDDAKRPRKTAPRDTGAFDVAEVBAMHR---  182
Db     139  DVRLSPDCEPFRRVKKIPGCKCKEWCDE-----PKDR---TAVGPALAAVRLDPT  186
QY     183  -----NCTAVTSPMSGCTSGGIGVSTRISVNNVNCQPEQSRSLCNLRPCDVDIH  232
Db     187  FGPPPTMMRANCLVQYTEMASCKSTGCMGISTVYTDNDFRLKESRSLCMRPPCAIDE  246
QY     233  TLINAGKRCIAYVQPEASNNFTLAGISTRSYQPKYQVGMCDNRCCIPIKYSKTIIDVSEQC  292
Db     247  ENIKGKKCIIRTPRIAPVVFELSGCSTVXTYPAKFCGVCTDRCCTPHRTTLPLVEFGC  306
QY     293  PDGLGFSROYLIAMINACNCNLSCRNPNDIFADL---SSYDSES  331
Db     307  PDGEIMKKNMFIITCACHNCPGDNIDFESLVIKRMATGMA  348

```

RESULT 4  
 A35669  
 gene CytR61 protein precursor - mouse  
 C:Species: Mus musculus (house mouse)  
 C:Date: 28-Sep-1990 #sequence\_revision 18-Nov-1992 #text\_change 05-Nov-1999  
 C:Accession: A35669; I48319; S16446  
 R:O'Brien, T.P.; Yang, G.P.; Sanders, L.; Lau, L.F.  
 Mol. Cell. Biol. 10, 3569-3577, 1990  
 A:Title: Expression of cytR61, a growth factor-inducible immediate-early gene.  
 A:Reference number: A35669; MUID:90287146  
 A:Accession: A35669  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-379 <OMB>  
 A:Cross-references: GB:M32490; NID:g192909; PIDN:AAA37512.1; PID:g309206  
 A:Note: the authors translated the codon GAT for residue 337 as Gln  
 Nucleic Acids Res. 19, 3261-3267, 1991  
 A:Title: Promoter function and structure of the growth factor-inducible immediate early  
 A:Reference number: I48319; MUID:91288203  
 A:Accession: I48319  
 A:Status: translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-379 <RES>  
 A:Cross-references: EMBL:X56790; NID:g50632; PIDN:CAA40109.1; PID:g50633  
 A:Note: the authors did not translate the codon for residue 108  
 A:Note: the authors translated the codon GAT for residue 337 as Gln  
 C:Genetics:  
 A:Gene: cytR61  
 A:introns: 21/3; 93/1; 208/1; 279/3  
 C:Superfamily: von Willebrand factor type C repeat homology  
 A:99-166/Domains: von Willebrand factor type C repeat homology <WCV>

[illegible]

QY 184 -----CIATSPWSPSTSCGLGVSTRISNVNAOCPEQESR 220  
 A:Accession: S20078  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-351 <YOL>  
 A:Cross-references: EMBL:X59284; NID:963702; PIDN:CAA1975.1; PID:963703  
 C:Genetics:  
 A:Gene: NOV

Db 204 KRLPVEGTBRVLFNPLHAHGOACIVQTTSMGCSKSCGISTRTYTNDRPECLVKEFR 263  
 QY 221 LCLRPDCDDIHTLTKAGKCLAVYOPESAMNTLAGCISTRSYQPKYCGVCDNRCCIP 280  
 Db 264 ICEVRPCGPPVSSLRKGRKSKTKKSPFVRYTAGCSVKKYRPRYKSCSYDGRCCIP 323  
 QY 281 YKSTIDVSFOCPDGLGFSROVLMINACFCNLSCRPNPND-----IFADLESPD 329  
 Db 324 LQRTYKMRFCEDGEMFSKNVMIOSCKONCPHNEASFLYLSLENDIHKFRD 379

RESULT 5  
 A41428  
 CEF-10 protein precursor - chicken  
 C:Species: Gallus gallus (chicken)  
 C:Date: 03-Apr-1992 #sequence\_revision 03-Apr-1992 #text\_change 21-Jul-2000  
 C:Accession: A41428  
 R:Simmons, D.L.; Levy, D.B.; Yannoni, Y.; Erikson, R.L.  
 Proc. Natl. Acad. Sci. U.S.A. 86, 1178-1182, 1989  
 A:Title: Identification of a phorbol ester-repressible v-src-inducible gene.  
 A:Reference number: A41428; MUID:89145206  
 A:Accession: A41428  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-375 <SIM>  
 A:Cross-references: GB:J04496; NID:9211435; PIDN:AAA48661.1; PID:9211436

Query Match 39.1%; Score 758; DB 2; Length 375;  
 Best Local Similarity 38.9%; Pred. No. 4,2e-49;  
 Matches 142; Conservative 55; Mismatches 100; Indels 68; Gaps 9;

QY 17 CKWPCPPSPRCPPLGVSLLITDGCCKMCAQOLGDNCTEAAICDPHRLGYCDYSGDRP 76  
 Db 26 CPAYCCCPAAAPQCAPGVGLVPGCGCKKVCANQLNDCSRTQPCDHTKGLKCNF--GASP 84  
 QY 77 RVALGVC-AQVAVGVLDGVRNNGOSFQPNKYNCTCIDGAVGCTPLC-LVYRPRRLW 134  
 Db 85 AATNGICRAQSEGRPEYNSKITYONGESFQPNCKHOCTCIDGAVGCTPLCPOELSLPILG 144  
 QY 135 CPAPRRVSIPEHCCQEWCEBDKAPRRKTPARDTGAFAVGEAVNH----- 181  
 Db 145 CPSPLRVKYPGQCCCEWVC-DESK-----DALEELGFFSKFGLDASGEL 190  
 QY 182 -RN-----CIATSPWSPSTSCGLGVSTRISNVN 210  
 Db 191 TRNNELIATVKGGLKMLPYGSEPOSRAPFNPKCIYQTTSMGCSKTCGISTRYVNDN 250  
 QY 211 AQCWPOESRLCNLRPCDDVDIHTLTKAGKCLAVYOPESAMNTLAGCISTRSYQPKYCG 270  
 Db 251 PDKLLETRICREVRPCGPPSYASLKKKKCTKTKKSPFVRYTAGCSVKKYRPRYCG 310  
 QY 271 VCMNDNCIPIYKSTIDVSFOCPDGLGFSROVLMINACFCNLSCRPNPNDIFADLESPD 330  
 Db 311 SCVDGRCCTPOOTRTYKIRPCDDGETTKSVMIOSCKONCPHNEASFLYLSLENDIHKFRD 379  
 QY 331 SEIAN 335  
 Db 364 YRLVN 368

RESULT 6  
 S20078  
 NOV protein - chicken  
 C:Species: Gallus gallus (chicken)  
 C:Date: 19-Feb-1994 #sequence\_revision 10-Nov-1995 #text\_change 21-Jul-2000  
 C:Accession: S20078  
 R:Joliet, V.; Martinerie, C.; Dambrine, G.; Plassiat, G.; Brisac, M.; Crochet, J.; Perh  
 Mol. Cell. Biol. 12, 10-21, 1992  
 A:Title: Proviral rearrangements and overexpression of a new cellular gene (nov) in myel  
 A:Reference number: S20078; MUID:92107157

A:Accession: S20078  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-351 <YOL>  
 A:Cross-references: EMBL:X59284; NID:963702; PIDN:CAA1975.1; PID:963703  
 C:Genetics:  
 A:Gene: NOV

Query Match 38.4%; Score 743.5; DB 2; Length 351;  
 Best Local Similarity 41.2%; Pred. No. 4.7e-48;  
 Matches 135; Conservative 49; Mismatches 117; Indels 27; Gaps 7;

QY 9 DTSSRPQFCWKPC--ECPPSPRCPPLGVSLLITDGCCKMCAQOLGDNCTEAAICDPHRLGYC 66  
 Db 23 EVSGREACPRPGRCGRPAAPPCAPVAVLDGCGCLVCARGRSGCSPLLPDDESG 82  
 QY 67 LYCDYSGDRPRVIAIGCAQVAVGVLDGVRNNGOSFQPNKYNCTCIDGAVGCTPLC- 125  
 Db 83 LYCD-RPEDEGGAGICMVEGDNVCFDGMITRNGEYFQSKYQCTCRDGOICLPN 141  
 QY 126 LVYRPRRLWCPHRRVSIPEHCCQEWCEBDKAPRRKTPARD--TGAFDVG----- 175  
 Db 142 LGILLPGDCPPFRKIEVPGCEKWCNCD-----PRDEVLLGCFMAAAYRQEATL 191  
 QY 176 --EYEAHNRNCIATSPWSPSTSCGLGVSTRISNVNAOCPEQESRLCNLRPCDDVDIHT 233  
 Db 192 GIVSDSSANCIQDTTMSASCSSKCGMFGSTRNRRQCEMYQTRLCMMRPEDE-EP 250  
 QY 234 LTKAGKCLAVYOPESAMNTLAGCISTRSYQPKYCGVCDNRCCIPYKSTIDVSFOCP 293  
 Db 251 SDKKGKCIOTKSKMAYRREYKNTSVOTYKRYGCLCNDGRCTPHNTKTIOVERCP 310  
 QY 294 DGLGFSRQVLMINACFCNLSCRPNPDI 321  
 Db 311 OGKFLKPRMLINTVCVCHGNCPOSNAF 338

RESULT 7  
 I38069  
 gene novH protein - human  
 C:Species: Homo sapiens (man)  
 C:Date: 17-May-1996 #sequence\_revision 17-May-1996 #text\_change 05-Nov-1999  
 C:Accession: I38069  
 R:Martinerie, C.; Huff, V.; Joubert, I.; Badzioch, M.; Saunders, G.; Strong, L.; Perh  
 Oncogene 9, 2729-2732, 1994  
 A:Title: Structural analysis of the human nov proto-oncogene and expression in Wilms  
 A:Reference number: I38069; MUID:94336229  
 A:Accession: I38069  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-357 <RES>  
 A:Cross-references: EMBL:X78351; NID:9587422; PIDN:CAA5146.1; PID:9825696  
 C:Genetics:  
 A:Gene: novH  
 A:Inserts: 28/3; 104/1; 188/1; 259/3  
 C:Superfamily: thrombospondin type 1 repeat homology  
 F:203-250/Domain: thrombospondin type 1 repeat homology <THRI>

Query Match 37.5%; Score 726.5; DB 2; Length 357;  
 Best Local Similarity 42.0%; Pred. No. 8.6e-47;  
 Matches 137; Conservative 46; Mismatches 116; Indels 27; Gaps 8;

QY 10 TSSRPQFCWKPC--ECPPSPRCPPLGVSLLITDGCCKMCAQOLGDNCTEAAICDPHRLGYC 69  
 Db 32 TORCPCPC--PGRCAPRPPPCAPVAVLDGCGCLVCARGRSGCSPLLPDDESG 89  
 QY 70 DYSGDRPRVIAIGCAQVAVGVLDGVRNNGOSFQPNKYNCTCIDGAVGCTPLC-LRV 128  
 Db 90 DRSAD-PSNOTGICTAIEGNCVDFDGIYRSGEYFQSKYQCTCRDGOICLPN 148  
 QY 129 RPRRLWCPHRRVSIPEHCCQEWVC--EDD-----AKRPRKTPARDTGAFAVNG-E 176



D0 Db 149 LLPEFNCAPPRKVEVPGECCEKWICGDPDEEDSLGLTLAARPEAT-----LGEV 198  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
OY 177 VEAHRNRINATSPWSPCSTSGIGSVSRISNVNAQCMPDPSRLCNRPDVD-INTLI 235  
| | | | | : | : | : | : | : | : | : | : | : | : | : | : | :  
Db 199 VSDSVCNCTEOTTEKTAKSKSCGMGFSTRVTNNRHCMLQTRLCMWRPCQEPEOPTD 258  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
OY 236 KAGKKCLAVYOPEASMNFTLAGCISTRSYOPRCGCVDNDNCRCCIPYKSRTIDVSFCPDG 295  
| | | | | : | : | : | : | : | : | : | : | : | : | : | : | :  
Db 259 KKGGKCILTKFKSKLSIAHLDQFNKNCTSLLHTYIKPFGCVCSIDGCCPHHNKTIQAERQCSBG 318  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
OY 296 LGFSROYLVMTINACFCNLSCRPNDFI 321  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
Db 319 QIVKKPVAVIGTCCTCHTNGCPNKNEAF 344  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
  
RESULT 8  
A43932  
mucin 2 precursor, intestinal - human (fragments)  
N: Alternate names: mucin SMOC-41  
C: Species: Homo sapiens (man)  
C: Date: 10-Mar-1993 #sequence revision 12-Apr-1996 #text change 05-Nov-1999  
C: Accession: A49963; A45106; B45106; A43932; B33532; A61257; PQ0328; PQ0329  
R: Gum J., J.R.; Hicks, J.W.; Toribara, N.W.; Siddiki, B.; Kim, Y.S.  
J: Biol. Chem. 265, 2440-2446, 1994  
A>Title: Molecular cloning of human intestinal mucin (MU2) cDNA. Identification of the  
A: Reference number: A49963; MUID: 94132002  
A: Accession: A49963  
A:Molecule type: mRNA  
A: Residues: 1-639 <GU1>  
A:Cross-references: GB:I21998  
R: Gum J., J.R.; Hicks, J.W.; Toribara, N.W.; Rothe, E.M.; Lagace, R.E.; Kim, Y.S.  
J: Biol. Chem. 267, 21375-21383, 1992  
A>Title: The human MU2 intestinal mucin has cysteine-rich subdomains located both upstr  
A: Reference number: A45106; MUID: 93016075  
A: Accession: A45106  
A>Status: not compared with conceptual translation  
A:Molecule type: mRNA  
A: Residues: 626-1895 <GU2>  
A:Cross-references: GB:M4131; NID:g186395; PIDN:AAA59163.1; PID:g186396  
A>Note: sequence extracted from NCBI backbone (NCBIP:116706)  
A: Accession: B45106  
A>Status: not compared with conceptual translation  
A:Molecule type: mRNA  
A: Residues: 2037-3020 <GU3>  
A:Cross-references: GB:M4132; NID:g186397; PIDN:AAA59164.1; PID:g186398  
A: Experimental source: colon  
A: Note: sequence extracted from NCBI backbone (NCBIP:116698)  
R: Toribara, N.W.; Gum J., J.R.; Culhane, P.J.; Lagace, R.E.; Hicks, J.W.; Petersen, G.N.  
J: Clin. Invest. 88, 1005-1013, 1991  
A>Title: MUC-2 human small intestinal mucin gene structure. Repeated arrays and polymor  
A: Reference number: A43932; MUID: 91356717  
A: Accession: A43932  
A:Molecule type: DNA  
A: Residues: 1343-1350, 'L', 1352-1411, 'S', 1413-1448, 'P', 1450-1503, 'T', 1505-1915 <TOR>  
A:Cross-references: GB:M74027; NID:g188863; PIDN:AAA59875.1; PID:g188864  
A: Note: sequence inconsistent with the nucleotide translation  
A: Note: sequence extracted from NCBI backbone (NCBIN:55749, NCBIP:55750)  
R: Gum, J.R.; Byrd, J.C.; Hicks, J.W.; Toribara, N.W.; Lamport, D.T.A.; Kim, Y.S.  
J: Biol. Chem. 264, 6480-6487, 1989  
A>Title: Molecular cloning of human intestinal mucin CDNA. Sequence analysis and eviden  
A: Reference number: A33532; MUID: 89197956  
A: Accession: B33532  
A:Molecule type: mRNA  
A: Residues: 1916-2193 <GU4>  
A:Cross-references: GB:M22405; NID:g188873; PIDN:AAA63534.1; PID:g188874  
A: Experimental source: intestine  
R: Janjy, B.H.; Gallup, M.W.; Yan, P.S.; Gum, J.R.; Kim, Y.S.; Baasbaum, C.B.  
J: Clin. Invest. 87, 77-82, 1991  
A>Title: Human bronchus and intestine express the same mucin gene.  
A: Reference number: A61257; MUID: 91086481  
A: Accession: A61257  
A>Status: not compared with conceptual translation

A:Molecule type: mRNA  
A:Residues: 'T',1925-1948,'TTS',1952-1954 <JAN>  
A:Experimental source: bronchus  
R:Xu, G., Huan, L., Khatami, I., Saitan, U.S., McCool, D.; Wang, D.; Jones, C.; Forstn  
Biochem. Biophys. Res. Commun. 183, 821-828, 1992  
A>Title: Human intestinal mucin-like protein (MUP) is homologous with rat MUP in the  
A:Reference number: P00328; MUID:92198477  
A:Accession: P00328  
A:Molecule type: mRNA  
A:Residues: 2328-2468 <XUG>  
A:Cross-references: GB:M86523  
A:Experimental source: small intestine  
A:Accession: P00329  
A:Molecule type: protein  
A:Residues: 2328-2342,'K',2344-2354 <XUG>  
C:Genetics:  
A:Gene: GDB:MUC2  
A:Cross-references: GDB:1120203; OMIM:158370  
A:Map position: 11p15.5-11p15.5  
C:Superfamily: von Willebrand factor; von Willebrand factor type A repeat homology; v  
C:Keywords: glycoprotein; intestine; tandem repeat  
F:2766-2834/Domains: von Willebrand factor type C repeat homology <WVC>

Query Match 7.88; Score 151.5; DB 2; Length 3020;  
Best Local Similarity 20.58; Pred. No. 0.0018;  
Matches 99; Conservative 40; Mismatches 156; Indels 189; Gaps 26;

QY 1 DPTAPLEDETSRRPOFCWPCCEPPSP--PRCPGLVSLITGCG-----ECCK----- 45  
|||:|||||  
Db 2527 DCTSPFLCOLIKDLSLFAOCHALVPQHYYDAC-----VFPSCTMPGSLSEACASLGAYAA 2580  
|||:|||||

QY 46 MCAQOL---GDNCTEAII---CDPHR-----GLYCDY 71  
:|||:|||||

Db 2581 ICAAOONICLDLRNRHTHGACTVECPSHREYACGPAREPPTCKSSSOONNTVLVAGCFPE 2640  
|||:|||||

QY 72 SGDRPRAIIGCAOVGVGCY-LDGVRNNQSGFPQNKRYKCTCIDAVG--CTR----- 123  
|||:|||||

Db 2641 G--TMNATAPGFDCVKTGCGVGPDNVPRFGHEFEFDCK-NCVCLGGSGILIQPKRCQ 2697  
|||:|||||

QY 124 -----LCILRVRP-----PRLMCP--HPRRVSIPIHCQ 148  
|||:|||||

Db 2698 KPVIHCVEDGYTLATEVPAPATCCNIIVCYKNSTLSCKRKPSV-CPLGEFVKSMVPGRC 2756  
|||:|||||

QY 149 EOWVCEDDAKKPRKTAIPDTGA-----FDVAGEVEAMHR--NCLAIYSPPWPCSTSC 198  
|||:|||||

Db 2757 PFYECESKGVCVRGNAEPQSPPYSSKCQCOCYCOTDKDNNTLLINVIACTH--VPCNTSC 2814  
|||:|||||

QY 199 GLGVSTRISNNACOMPQESRLCNLRCPDVDIHTLIKAGK-----NCLAIYQ 246  
|||:|||||

Db 2815 SPGR--ELMEAPGEGCKCKCEOTHCIKRPD-NOHVILPQDFKSDPKNKNCTFFSCVKIHN 2871  
|||:|||||

QY 247 PEASN-----NFTLAGCI-STRSYOPKYC----- 269  
|||:|||||

Db 2872 QLISVSNTICPNPDASICIGSITFMNGCKCKTCPRNETRVPCSVTPVTTEVSYAQT 2931  
|||:|||||

QY 270 -----GYC-----MDN--RCCTPYKSTIDIVSQCPDGLFSQVQLMIN 306  
|||:|||||

Db 2932 KTVILMNHCSSCGTFVMYSANAQAOLDHSCSCCKKEKTSOREVLSCPWGSLTHTYTHIE 2991  
|||:|||||

QY 307 ACFC 310  
|||:|||||

Db 2992 SCOC 2995

RESULT 9  
A46019  
Notch-1 protein - mouse  
N:Alternate names: notch protein  
C:Species: Mus musculus (house mouse)  
C>Date: 22-Sep-1993 #sequence\_revision 18-Nov-1994 #text\_change 20-Sep-1999  
C:Accession: A46019; S215144  
C:Idel Amo, F.F.; Genotro-Maguire, M.; Swiatek, P.J.; Jenkins, N.A.; Copeland, N.G.; G

Genomics 15, 259-264, 1993  
A:Title: Cloning, analysis, and chromosomal localization of Notch-1, a mouse homolog of  
A:Reference number: A46019; MUID:93194170  
A:Accession: A46019  
A:Status: not compared with conceptual translation  
A:Molecule type: nucleic acid  
A:Residues: 1-2531 <DEL>  
A:Cross-references: GB:21886; GB:S47228; NID:q288502; PIDN:CAA77941.1; PID:q288503  
A:Note: sequence extracted from NCBI Backbone (NCBI:127318)  
R:Francisco del Amo, F.; Smith, D.E.; Swatek, P.J.; Gendron-Maguire, M.; Greenspan, R.J.;  
submitted to the EMBL Data Library, April 1992  
A:Description: The expression pattern of Notch, a mouse homolog of Drosophila Notch, suggests  
A:Reference number: 525144  
A:Accession: 525144  
A:Molecule type: mRNA  
A:Residues: 1551-2108, 'Q', 2110-2114, 'ALP', 2118-2170 <FRA>  
A:Cross-references: EMBL:Z11886  
C:Genetics:  
A:Gene: notch-1  
A:Map position: 2  
A:Note: proximal region of chromosome 2  
C:Superfamily: unassigned ankyrin repeat proteins; ankyrin repeat homology; EGF homology  
F:106-138/Domain: EGF homology <EGF1>  
F:144-175/Domain: EGF homology <EG01>  
F:222-254/Domain: EGF homology <EG02>  
F:261-292/Domain: EGF homology <EG02>  
F:339-370/Domain: EGF homology <EG03>  
F:416-449/Domain: EGF homology <EGF3>  
F:456-487/Domain: EGF homology <EG04>  
F:494-525/Domain: EGF homology <EG05>  
F:532-563/Domain: EGF homology <EG06>  
F:607-638/Domain: EGF homology <EG07>  
F:682-713/Domain: EGF homology <EG08>  
F:757-788/Domain: EGF homology <EG09>  
F:795-826/Domain: EGF homology <EG10>  
F:873-904/Domain: EGF homology <EG11>  
F:911-942/Domain: EGF homology <EG12>  
F:949-980/Domain: EGF homology <EG13>  
F:987-1018/Domain: EGF homology <EG14>  
F:1025-1056/Domain: EGF homology <EG15>  
F:1063-1094/Domain: EGF homology <EG16>  
F:1149-1180/Domain: EGF homology <EG17>  
F:1187-1218/Domain: EGF homology <EG18>  
F:1233-1264/Domain: EGF homology <EGF4>  
F:1352-1383/Domain: EGF homology <EG19>  
F:1391-1425/Domain: EGF homology <EGF>  
F:1917-1948/Domain: ankyrin repeat homology <AN1>  
F:1949-1981/Domain: ankyrin repeat homology <AN2>  
F:1983-2015/Domain: ankyrin repeat homology <AN3>  
F:2016-2048/Domain: ankyrin repeat homology <AN4>  
F:2049-2081/Domain: ankyrin repeat homology <AN5>

Oy 169 GAPDAVGVEVMHNRNCIATVTSFMSPCS---ISCLGVS-----TRISNNVAQCWP- 215  
 Db 403 G-----YTPGP---ACSDVDVDCDGLGANRCENHACKCLNTLIGSEPCCLQG 443  
 Oy 216 -----ESESRLCNLRPDDVDIHTLIRAGKKCLAVYQPEASMNFTLACGISTRSPYKRC 269  
 Db 444 YTGPGCELDVNEGCSNPGQNDATCLDQIG-----FO-----CLCMPEYEGVYC 487  
 Oy 270 GV-----CMDNRCIPYKSKTIDVSFQCPDGLGFSRQVLMINACFNLCRN 316  
 Db 488 EINTDECASSPCLNHGHCN---DKIHFPCCCPKGFNGHLCQYDVDEC-ASTPCKN 539  
 RESULT 10  
 VMHU  
 von Willebrand factor precursor - human  
 C:Species: Homo sapiens (man)  
 C:Date: 04-Dec-1986 #sequence.revision 30-Jun-1993 #text.change 22-Jun-1999  
 C:Accession: A34480; S02377; A37139; S23676; A25469; A23366; S23618; S23645;  
 R:Mancuso, D.J.; Tulay, E.A.; Westfield, L.A.; Worral, N.K.; Shelton-Inloes, B.B.; S  
 J. Biol. Chem. 264, 19514-19527, 1989  
 A:Title: Structure of the gene for human von Willebrand factor.  
 A:Accession: A34480; MUID:90062044  
 A:Molecule type: DNA  
 A:Residues: 1-2813 <MAN>  
 A:Cross-references: EMBL:M25864  
 R:Bonthron, D.; Orkin, S.H.  
 Eur. J. Biochem. 171, 51-57, 1988  
 A:Title: The human von Willebrand factor gene. Structure of the 5' region.  
 A:Reference number: S02377; MUID:8811704  
 A:Accession: S02377  
 A:Molecule type: DNA  
 A:Residues: 1-177 <BO2>  
 A:Cross-references: EMBL:X06828  
 R:Mancuso, D.J.; Tulay, E.A.; Westfield, L.A.; Lester-Mancuso, T.L.; Le Beau, M.M.; S  
 Biochemistry 30, 253-269, 1991  
 A:Title: Human von Willebrand factor gene and pseudogene: structural analysis and dif  
 A:Reference number: A37139; MUID:91105089  
 A:Accession: A37139  
 A:Molecule type: DNA  
 A:Residues: 990-1947 <MAD>  
 A:Cross-references: GB:M06075; NID:g340357; PIDN:AAA61295.1; PID:g553810  
 A:Note: the authors translated the codon CGC for residue 156 as Gln  
 R:Collins, C.J.; Underdahl, J.P.; Levene, R.B.; Ravera, C.P.; Morin, M.J.; Dombalaja  
 Proc. Natl. Acad. Sci. U.S.A. 84, 4393-4397, 1987  
 A:Title: Molecular cloning of the human gene for von Willebrand factor and identifica  
 A:Reference number: S23676; MUID:87260814  
 A:Accession: S23676  
 A:Molecule type: DNA  
 A:Residues: 2731-2813 <COL>  
 A:Cross-references: EMBL:M16945  
 R:Bonthron, D.; Orr, E.C.; Mitsock, L.M.; Ginsburg, D.; Handin, R.I.; Orkin, S.H.  
 Nucleic Acids Res. 14, 7125-7127, 1986  
 A:Title: Nucleotide sequence of pre-pro-von Willebrand factor cDNA.  
 A:Reference number: A25298; MUID:87016349  
 A:Accession: A25298  
 A:Molecule type: mRNA  
 A:Residues: 1-470, 'V', 472-2813 <BON>  
 A:Cross-references: EMBL:X04385  
 R:Verweij, C.L.; Diergaarde, P.J.; Hart, M.; Pannekoek, H.  
 EMBO J. 5, 1839-1847, 1986  
 A:Title: Length of von Willebrand factor (vWF) cDNA encodes a highly repetitive pro  
 A:Reference number: A91044; MUID:87004550  
 A:Accession: A25469  
 A:Molecule type: mRNA  
 A:Residues: 1-470, 'V', 472-483, 'R', 485-1022, 'K', 1024-1025, 'E', 1027-1400 <VER>  
 A:Cross-references: EMBL:X04146  
 A:Note: this sequence has been revised in reference A91056  
 R:Verweij, C.L.; Diergaarde, P.J.; Hart, M.; Pannekoek, H.  
 EMBO J. 5, 3074, 1986  
 A:Reference number: A91056  
 A:Accession: A25366

A: Molecule type: mRNA  
 A: Residues: 1021-1030 <VE2>  
 A: Note: this is a revision to the sequence from reference A91044  
 R: Shelton-Inloes, B.B.; Broze Jr., G.J.; Miletich, J.P.; Sadler, J.E.  
 Biochem. Biophys. Res. Commun. 144, 657-665, 1987  
 A: Title: Evolution of human von Willebrand factor: cDNA sequence polymorphisms, repeated  
 A: Reference number: S23618; MUID: 871213253  
 A: Accession: S23618  
 A: Molecule type: mRNA  
 A: Residues: 1-120 <SR2>  
 A: Cross-references: EMBL:M17588; NID: g799330; PIDN: AAA65940.1; PID: g340316  
 A: Accession: S23645  
 A: Molecule type: protein  
 A: Residues: 23-56 <SH3>  
 R: Sadler, J.E.; Shelton-Inloes, B.B.; Sorace, J.M.; Harlan, J.M.; Tiliak, K.; Davie, E.W.  
 Proc. Natl. Acad. Sci. U.S.A. 82, 6394-6398, 1985  
 A: Title: Cloning and characterization of two cDNAs coding for human von Willebrand factor  
 A: Reference number: A94060; MUID: 86016708  
 A: Accession: A94060  
 A: Molecule type: mRNA  
 A: Residues: 'WA', '739', 'C', '744'-769, 'H', '771-788, 'A', '790-803, 'S', '805-873, 1289-1471, 'D', '1473-  
 A: Note: the authors translated the codon TCG for residue 2168 as Cys  
 R: Shelton-Inloes, B.B.; Tiliak, K.; Sadler, J.E.  
 Biochemistry 25, 3164-3171, 1986  
 A: Title: cDNA sequences for human von Willebrand factor reveal five types of repeated dc  
 A: Reference number: A90504; MUID: 86269894  
 A: Accession: A90504  
 A: Molecule type: mRNA  
 A: Residues: 781-788, 'A', '790-1424 <SHE>  
 A: Note: 852-Gln, 857-Asp, and 1381-Thr were also found  
 R: Ginsburg, D.; Handin, R.I.; Bonthron, D.T.; Donlon, T.A.; Bruns, G.A.P.; Latt, S.A.; C  
 Science 228, 1401-1406, 1985  
 A: Title: Human von Willebrand factor (VWF): Isolation of complementary DNA (cDNA) clones  
 A: Reference number: A44178; MUID: 85244588  
 A: Accession: A44178  
 A: Molecule type: mRNA  
 A: Residues: 2621-2813 <GIN>  
 A: Cross-references: EMBL: K03028; NID: g340308; PIDN: AAA61293.1; PID: g340309  
 R: Verweij, C.L.; de Vries, C.J.M.; Distel, B.; van Zonneveld, A.D.; van Kessel, A.G.; va  
 Nucleic Acids Res. 13, 4699-4717, 1985  
 A: Title: Construction of cDNA coding for human von Willebrand factor using antibody prot  
 A: Reference number: S07363; MUID: 85269803  
 A: Accession: S07363  
 A: Molecule type: mRNA  
 A: Residues: 2731-2813 <VE3>  
 A: Cross-references: EMBL: X02672; NID: g37939; PIDN: CAA26503.1; PID: g37940  
 R: Lynch, D.C.; Zimmerman, T.S.; Collins, C.J.; Brown, M.; Morin, M.J.; Ling, E.H.; Lavin  
 Cell 41, 49-56, 1985  
 A: Title: Molecular cloning of cDNA for human von Willebrand factor: authentication by a  
 A: Reference number: S23678; MUID: 85201887  
 A: Accession: S23678  
 A: Molecule type: mRNA  
 A: Residues: 2731-2813 <LYN>  
 A: Cross-references: EMBL: K03028  
 R: Tiliak, K.; Kumar, S.; Takio, K.; Ericsson, L.H.; Wade, R.D.; Ashida, K.; Walsh, K.A.;  
 Biochemistry 25, 3171-3184, 1986  
 A: Title: Amino acid sequences of human von Willebrand factor.  
 A: Reference number: A90505; MUID: 86269895  
 A: Accession: A90505  
 A: Molecule type: protein  
 A: Residues: 764-788, 'A', '790-1471, 'D', '1473-2813 <TIT>  
 A: Note: 789-Thr was also found  
 R: Chopok, M.W.; Girma, J.P.; Fujikawa, K.; Davie, E.W.; Tiliak, K.  
 Biochemistry 25, 3146-3155, 1986  
 A: Title: Human von Willebrand factor: a multivalent protein composed of identical subunit  
 A: Reference number: A23464; MUID: 86269892  
 A: Accession: A23464  
 A: Molecule type: protein  
 A: Residues: 764-773, 2803-2813 <CHO>  
 R: Ident, J.A.; Berkowitz, S.D.; Ware, J.; Kasper, C.R.; Ruggeri, Z.M.  
 Proc. Natl. Acad. Sci. U.S.A. 87, 6306-6310, 1990  
 A: Title: Identification of a cleavage site directing the immunochemical detection of mol  
 A: Reference number: A36013; MUID: 90349604

A: Accession: A36013  
 A: Molecule type: protein  
 A: Residues: 1606-1617 <DEN>  
 R: Fay, P.J.; Kawai, Y.; Wagner, D.D.; Ginsburg, D.; Bonthron, D.; Ohlsson-Wilhelm, B.  
 Science 232, 995-998, 1986  
 A: Title: Propolyptide of von Willebrand factor circulates in blood and is identical  
 A: Reference number: A60913; MUID: 86208144  
 A: Accession: A60913  
 A: Molecule type: protein  
 A: Residues: 576-590 <RAY>  
 A: Accession: GDB: VWF  
 A: Cross-references: GDB: 119125; OMIM: 193400  
 A: Map position: 12p13.3-12p13.2  
 A: Intons: 19/1; 74/1; 108/2; 178/1; 219/3; 292/1; 333/1; 370/2; 386/1; 431/3; 478/1;  
 5/1; 1724/1; 1771/1; 1819/1; 1888/3; 1948/1; 2021/3; 2086/1; 2200/1; 2266/3;  
 C: Keywords: blood coagulation; cell binding; connective tissue; disulfide bond; dupl1  
 F: 1-22/Domain: signal sequence \*status predicted <SIC>  
 F: 23-763/Domain: von Willebrand antigen II \*status predicted <MA1>  
 F: 34-386/Domain: type D repeat 1 <DD1>  
 F: 387-745/Domain: type D repeat 2 <DD2>  
 F: 768-700/Region: cell attachment (R-G-D) motif  
 F: 764-2813/Region: von Willebrand factor \*status predicted <MA2>  
 F: 784-865/Domain: D' <DD2>  
 F: 788-833, 2216-2261/Region: duplication  
 F: 826-853, 2400-2515, 2544-2662/Region: duplication  
 F: 842-1130, 1934-2203/Region: duplication  
 F: 866-1241/Domain: type D repeat 3 <DD3>  
 F: 1275-1443/Domain: von Willebrand factor type A repeat homology <VWA1>  
 F: 1436-1654/Domain: von Willebrand factor type A repeat homology <VWA2>  
 F: 1689-1854/Domain: von Willebrand factor type A repeat homology <VWA3>  
 F: 1947-2295/Domain: type B repeat 4 <BD4>  
 F: 2296-2330/Domain: type B repeat 4 <BD4>  
 F: 2340-2365/Domain: type B repeat 2 <VB2>  
 F: 2375-2399/Domain: type B repeat 3 <VB3>  
 F: 2430-2497/Domain: von Willebrand factor type C repeat homology <VWC1>  
 F: 2507-2509/Region: cell attachment (R-G-D) motif  
 F: 2581-2647/Region: von Willebrand factor type C repeat homology <VWC2>  
 F: 2817, 1231, 1515, 1574, 2223, 2290, 2357, 2400, 2546, 2585, 2790/Binding site: carbohydrate (A  
 F: 1147/Binding site: carbohydrate (Asn) (covalent) \*status atypical  
 F: 1248, 1255, 1256, 1468, 1477, 1487, 1679, 2298/Binding site: carbohydrate (Thr) (covalent)  
 F: 1263, 1486/Binding site: carbohydrate (Ser) (covalent) \*status experimental

Query Match 7.7% Score 150; DB 1: Length 2813;  
 Best local similarity 22.9% Pred. No. 0.0022;  
 Matches 93; Conservative 43; Mismatches 112; Indels 158; Gaps 31;

QY	14	POFCMPCPCPPSPPPRCPLGSLITDGE--CCRMK--AAQAGNCTEAAICDPHRLG	68
DB	2196	PDFC--AMSCP-----SLVYNHCBBGCRHCDGVSSCGDHPSCGCPDPKVM	2244
QY	69	---CDYSGRRPRTAIGCAQVNVGCVLDGVYNNQSGOPN---CKYVCTCIDG-AVGC	121
DB	2245	EGSC-----VPEEA---CTQCGE---DGVOHGFLEAWVPDHPQGI-CTCLSGRKVNC	2291
QY	122	T-----PLC-----LRVP-----PRLMC-----PHPR-----VSP	144
DB	2292	TTQPCPTAKAPTCGICFVARLQNMADCCRETCVCDPVSCDLPVPHCEKGLDPLTNP	2351
QY	145	GHSCBQWNC---EDDAKR--PRKTAP-----RDTGAFDAVEVNAHNRCLAVTSPWS	193
DB	2352	GBCRNPFCAKREBECKRVRSPSPCRHRLPTLRKTCOCQ---EYEC-ACNCVNST-----	2402
QY	194	CTSGCLG--VSTRISN---VNAQCPDESRCLNRP-----CVDYHTLTIK	236
DB	2403	--VSCPGLYASTAINDCGCTTTTCLPD---KVCHVRSITVPVCGFWERGCYV-----	2450
QY	223	ACKKCLAVYQPARSNFTLAGCI-----STRS-----YORKYGVGCDMNRGCIPIYS-	283
DB	2451	---CTCTDMDNAVGLRVAQCSQKPCEDSCRSRGTFTYLVHBECCGRCLPSACEVVTGSP	2506

QY 284 -----KITDVSFOCPDGLGFSROVLMINACFCNLSCNPNDIF 321  
 DB 2507 RGDSSQSKSVSGQMSAPE-----NPLINCVAVKKEVF 2541

# RESULT 11

A39804  
 thrombospondin precursor - chicken

C:Species: Gallus gallus (chicken)

C>Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999

C:Accession: A39804

R:Lawler, J.; Duquette, M.; Ferro, P.

J. Biol. Chem. 266, 8039-8043, 1991

A:Title: Cloning and sequencing of chicken thrombospondin.

A:Reference number: A39804, MUID:91217026

A:Accession: A39804

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-1178 <LAV>

C:Superfamily: GB:M60853, NID:q212763, PIDN:AAA51437.1, PID:q212764

C:Superfamily: Thrombospondin 1; EGF homology; thrombospondin type 1 repeat homology; von

F:325-383/Domain: von Willebrand factor type C repeat homology <VWC>

F:386-437/Domain: thrombospondin type 1 repeat homology <THR1>

F:442-498/Domain: thrombospondin type 1 repeat homology <THR2>

F:499-555/Domain: thrombospondin type 1 repeat homology <THR3>

F:658-697/Domain: EGF homology <EGF>

Query Match 7.6%; Score 147.5; DB 1; Length 1178;  
 Best Local Similarity 21.8%; Pred. No. 0.0016;  
 Matches 86; Conservative 38; Mismatches 121; Indels 149; Gaps 23;

QY 8 EDTSSRPQCKMPCEPPSPRCPLGVSLITDGCCECKMCAQOLDGNCTEAA----- 59  
 DB 263 EKTEKAEFCEDRCE-----ELG-TMFE-LTGRLVNNINLADMIQVSEENQIMTE 312  
 QY 60 ICPHRLGLYCDYSGDRPRAIGVCAQVVGCVLDGVRVYNNNGSFQPNCKYNCCTIDGAV 119  
 DB 313 LIGPNKTLKQ-----SVCWQ-----DGRVFPDSESIYDSCCTKCTCQDSKI 354  
 QY 120 GCTPLCLRVPRPRLMCPHRRVSIIPHGC-----EQW-----VCED--- 155  
 DB 355 ----VCHQIICPPYSCADPS--FIESECCPVCHSDSESGMSPWMTKCSVYCGSGTQ 408  
 QY 156 ----DAKPRKTPA-----RDGAFDAVGEVEAHNRNCIAYTSPWSP 193  
 DB 409 MRGRSCDVTRSACTGPHIQTIRMCSFKKCDHRIRQDG-----GW-----SHMSPWSS 454  
 QY 194 CSTRSCGVSTRISNVNAQWPE-----QESRLCNLRPCDDVDIHTLIRAGKKCLA 243  
 DB 455 CSVTCGNGITRIRLCLNSPI-PQMGKNCVNGRETEKCKRACPVA-----GQ---WG 504  
 QY 244 VYQPEASMNFTLACISTSTRY-----QPKYGVCMNRCICPYKSKTIDVSFOCP-DGLG 297  
 DB 505 PMSPMWACITVTCGGGIFRERLCLNSPEPYG---KPCVGDITKQHMCKMRCPIDG-- 558  
 QY 298 FSRQVLMINACFCNLSCRPNDIFADLESYDPDS 331  
 DB 559 ----CLSNPCFPGAEC-----NSYPDGS 577

# RESULT 12

A57534  
 mucin 5AC (clone L31) - human (fragment)

C:Species: Homo sapiens (man)

C>Date: 08-Feb-1996 #sequence\_revision 08-Feb-1996 #text\_change 20-Apr-2000

C:Accession: A57534

R:Resutflaur, T.; Roche, F.; Hill, A.S.; Lacasa, M.; Fox, M.; Swallow, D.M.; Zweibaum, A.

J. Biol. Chem. 270, 13665-13673, 1995

A:Title: Characterization of a mucin cDNA clone isolated from HT-29 mucus-secreting cell

A:Reference number: A57534, MUID:95293957

A:Accession: A57534

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-1042 <LES>

A:Cross-references: GB:248314, NID:q1052607, PIDN:CAA88307.1, PID:q1052608

C:Genetics:

A:Gene: GDB:MUC5AC

A:Cross-references: GDB:454136; OMIM:158373

A:Map position: 11p15.5-11p15.5

C:Superfamily: von Willebrand factor type C repeat homology

F:678-746/Domain: von Willebrand factor type C repeat homology <VWC>

Query Match 7.5%; Score 145.5; DB 2; Length 1042;  
 Best Local Similarity 20.9%; Pred. No. 0.002;  
 Matches 84; Conservative 29; Mismatches 137; Indels 151; Gaps 23;

QY 13 RPOGCKWPCCEPPSPRCPL---GVSILITDGCCEC-----KMCADQOLDGNCTEAA 59  
 DB 607 RPKLC-----PLRPACPLPFGVPPAPADQAGCCQYSCACNTSRCPAPVG--CPGA 657  
 QY 60 ICPHRLGLYCDYSGDRPRAIGVCAQVVG---CVLDGVRVYNNNGSFQPNCKYNCCTC-I 115  
 DB 658 ----RAI-----PTYDGAACCPVQNCSTVCSINGLYOPGAVVSSLCETCRCEL 704  
 QY 116 DG-----AVGC-TPICLRVPRPRLMCPHRRVSI-----PGHCCEQWCEDDAKRPRK 162  
 DB 705 PGGPPSAFVYVSCETQIC-----NTHCP---VGFEYQEOSGCC--GTGVQVA----- 747  
 QY 163 TAPRDGAFDAVGEVEAHNRNCIAYTS-----PMSPCSTSCGLVSTRISVNA 211  
 DB 748 ----CVTNTSKSPAHLFPYGPETMSDAGNHC-----VTH 776  
 QY 212 QCPPEQE-----SRICNLPRCDVDIHTLIKAG-----KKLAVYQPEAS 250  
 DB 777 QCEKHQDGLVYVTKKACRPLSCSLDEARMSKDCRCRPLRPYPYONQSTAVY--HRS 834  
 QY 251 MNFTLACISSTRYQPYKC-GVCMND-----RCCIPYKSKTIDVSFOCPD 294  
 DB 835 LIIDQGCSSSEPRVRLAYCRKNGCDSSMYSLEGNVYEHRCQCCQELRTSLRNTYLCTD 894  
 QY 295 GLGFSROVLMINACFC-NLSCRPNDIFADLESYDPFSEIA 334  
 DB 895 GSSRAFSYTEVEECGCMGRCPAPDQHSSEAPERSOEA 935

# RESULT 13

A42112  
 mucin-like peptide MLP 2677 - rat

C:Species: Rattus norvegicus (Norway rat)

C>Date: 04-Mar-1993 #sequence\_revision 18-Nov-1994 #text\_change 10-Oct-1997

C:Accession: A42112

R:Xu, G.; Huan, L.J.; Khatri, I.A.; Wang, D.; Bennick, A.; Fahlin, R.E.; Forstner, G.G.

J. Biol. Chem. 267, 5401-5407, 1992

A:Title: cDNA for the carboxyl-terminal region of a rat intestinal mucin-like peptide

A:Reference number: A42112, MUID:92184794

A:Accession: A42112

A:Status: preliminary

A:Molecule type: nucleic acid; protein

A:Residues: 1-837 <XU1>

A:Experimental source: intestine

A>Note: sequence inconsistent with the nucleotide translation

C:Superfamily: von Willebrand factor type C repeat homology

F:582-650/Domain: von Willebrand factor type C repeat homology <VWC>

Query Match 7.5%; Score 144.5; DB 2; Length 837;  
 Best Local Similarity 20.1%; Pred. No. 0.002;  
 Matches 84; Conservative 37; Mismatches 133; Indels 163; Gaps 25;

QY 23 CPSPS-PPRC---PLGVSLITDGC-----ECCQKCA-----QQLGDN- 54  
 DB 429 GPPSEPTQSSSPKNSLTLLVGGCFEGTTFKPAFYDVGVKICGCVGPDVVPREFGEHF 488

OY 55 -----CTEA-----ICDPR-----GLY-----CD-----70  
 DB 489 EFDKDCVCLBEGSGIVCPKCAKANGMLTTCEDDTYLVEADPDCKCNTTSCDKPKR 548  
 OY 71 YSGDR-----RYAIGCAQV-----GYGCVLDGVYNNQSGFQPCNKYKNCI 115  
 DB 549 CAERPSCLGFEVSEHVPKCCPVYSCVRKV-CYHENAIFYOGSPYKSKODCVCT 607  
 OY 116 DGAVCTPL-----CLRPRPRLMCPHP-RRVSIPGHCEQWVCECD--AKRPR-----161  
 DB 608 DSMDSSTOLNIVISCTHV-PCNISCSGSEFELVEPECCCK--CQOHTCIKRPQOYIL 664  
 OY 162 -----KTAIPRODIF-----DANGFY--EAMRNQIATSPMSF--CSTSCGL 200  
 DB 665 KPEIQRNPNDKCTFPCSKKINNOLISSVNTCPDFPDSDCVPSTIYMPNGCKCTC--722  
 OY 201 GVSTRISVNAOCWPEOESRLCNLRPC-----DVDITLIKAKKCLAVQPEASMF 253  
 DB 723 -----IHNPN-----NTVPCSAIPYKKEISTYNGCAK-----NISMF 754  
 OY 254 TIAGCTSTRSYQKYCGVCMNRCIPYKSKTIDVSFQCPDGLGFSROYLWIMNACFC 310  
 DB 755 CAGSGCTFAMYSQAQODLDHGCSCCRERTSVRMVSLDCCPDGSKLSHYTHIESCIC 811

## RESULT 14

anglogenes is inhibitor homolog - Caenorhabditis elegans  
 C:Species: Caenorhabditis elegans  
 C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 18-Feb-2000  
 C:Accession: T18855; T24653  
 R:McMurray, A.  
 Submitted to the EMBL Data Library, July 1995  
 A:Reference number: Z19031  
 A:Accession: T18856  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-1444 <N144>  
 A:Cross-references: EMBL:Z50004; PIDN:CAA90293.1; GSPDB:GN00028; CESP:CO2B4.1  
 R:McMurray, A.  
 Submitted to the EMBL Data Library, July 1995  
 A:Reference number: Z19917  
 A:Accession: T24653  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-1444 <N12>  
 A:Cross-references: EMBL:Z50006; PIDN:CAA90302.1; GSPDB:GN00028; CESP:CO2B4.1  
 A:Experimental source: clone T07C5  
 C:Genetics:  
 A:Gene: CESP:CO2B4.1  
 A:Map position: X  
 A:Introns: 25/3; 70/3; 96/3; 139/3; 187/1; 234/2; 282/3; 376/2; 422/2; 478/3; 509/3; 566

Query Match 7.4%; Score 144; DB 2; Length 1444;  
 Best Local Similarity 25.4%; Pred. No. 0.0035;  
 Matches 70; Conservative 20; Mismatches 66; Indels 120; Gaps 18;  
 OY 10 TSSROFC-----KWC-----ECPSPRRCPLGLVSLITDGECCCKMACQ 49  
 DB 1149 TSTRRFPCVDPYVQFCAGAILRDIIPCAPGSCSPSAG-----GMSLWSEWSSCKDC--1202  
 OY 50 QIGDN-----CTEAICDPHRLGYCD-YSGD-RPRYAIGVCAQVGVGCVLDGVRYN 99  
 DB 1203 --GDGHOIRNMCMGRP--IPSNRGATGCGYSFDQRP-----CYMNV-----1241  
 OY 100 NQSGFQPCNKYKNCI DGAVG-----CTPLCLVRPRRL-WCPHPRVSIPIGHC 147  
 DB 1242 -----CSDEKVDGKWTDMTAMSECTDYCRNGHRSRTFCANPKPSGGAQC 1287  
 OY 148 C-----EOWCEDAKRPRRTARPDGAFDAVGEVEAMHNCIAIYSPWPCSTSGCLGYS 203

DB 1288 TGSDFELNCFD-----PARCLRDCG-----W-----STWSDMTPCSASCGFCVQ 1328  
 OY 204 TRISVNAOCWPE-----CESRLCNLRPCD 228  
 DB 1329 TRDRSCSS--PEPRGGGSCSGLAHQTSICDLPAQD 1361

## RESULT 15

TSHDPI  
 thrombospondin 1 precursor - human  
 C:Species: Homo sapiens (man)  
 C:Date: 23-Aug-1987 #sequence\_revision 03-Aug-1995 #text\_change 13-Aug-1999  
 C:Accession: A26155; A34274; A30140; A25812; A05172; A42927  
 R:Lawler, J.; Hynes, R.O.  
 J. Cell Biol. 103, 1635-1648, 1986  
 A>Title: The structure of human thrombospondin, an adhesive glycoprotein with multiple  
 A:Accession number: A26155; MUID:87057617  
 A:Molecule type: mRNA  
 A:Residues: 1-1170 <LAW>  
 A:Cross-references: GB:X04665; NID:937137; PIDN:CAA28370.1; PID:937138  
 A>Note: parts of this sequence, including the amino end of the mature protein, were d  
 R:Labery, C.D.; Gierman, T.M.; Dixit, V.M.  
 J. Biol. Chem. 264, 11222-11227, 1989  
 A>Title: Characterization of the promoter region of the human thrombospondin gene. DN  
 A:Reference number: A34274; MUID:89291870  
 A:Accession: A34274  
 A:Molecule type: DNA  
 A:Residues: 1-166 <LAW>  
 A:Cross-references: GB:J04835  
 R:Hennessey, S.W.; Frazer, B.A.; Kim, D.D.; Deckwerth, T.L.; Baumgartel, D.M.; Rotwel  
 J. Cell Biol. 108, 729-736, 1989  
 A>Title: Complete thrombospondin mRNA sequence includes potential regulatory sites in  
 A:Reference number: A30140; MUID:89139590  
 A:Accession: A30140  
 A:Molecule type: mRNA  
 A:Residues: 1-83, 'A', 85-522, 'A', 524-1170 <HEN>  
 A:Cross-references: EMBL:X14787; NID:937464; PIDN:CAA32889.1; PID:937465  
 A>Note: parts of this sequence, including the amino end of the mature protein, were d  
 R:Kobayashi, S.; Eden-McCutchan, F.; Framson, P.; Bornstein, P.  
 Biochemistry 25, 8418-8425, 1986  
 A>Title: Partial amino acid sequence of human thrombospondin as determined by analysi  
 A:Reference number: A25812; MUID:87157592  
 A:Accession: A25812  
 A:Molecule type: mRNA  
 A:Residues: 1-83, 'A', 85-397 <KOB>  
 A:Cross-references: GB:M25631; NID:9538353; PIDN:AAA36741.1; PID:9538354  
 R:Dixit, V.M.; Hennessey, S.W.; Grant, G.A.; Rotwein, P.; Frazer, W.A.  
 Proc. Natl. Acad. Sci. U.S.A. 83, 5449-5453, 1986  
 A:Reference number: A05172; MUID:86287276  
 A:Accession: A05172  
 A:Molecule type: mRNA  
 A:Residues: 1-83, 'A', 85-374, 'RC', <DIX>  
 A:Cross-references: GB:M14326; NID:9340005; PIDN:AAA61237.1; PID:9553801  
 A>Note: parts of this sequence, including the amino end of the mature protein, were d  
 R:Sun, X.; Skorstengaard, K.; Mosher, D.F.  
 J. Cell Biol. 118, 693-701, 1992  
 A>Title: Disulfides modulate RGD-inhibitable cell adhesive activity of thrombospondin  
 A:Reference number: A42927; MUID:92348511  
 A:Accession: A42927  
 A:Molecule type: Protein  
 A:Residues: 987-1003 <SUN>  
 A>Note: Cys-992 is shown to have a free sulfhydryl  
 C:Genetics:  
 A:Gene: GDB:THBS1; TSP1; TSP  
 A:Cross-references: GDB:120438; OMIM:188060  
 A:Map position: 15q15-15q15  
 A:Introns: 23/1  
 A>Note: the list of introns may be incomplete  
 C:Complex: homotrimer, disulfide linked  
 C:Function:  
 A:Description: participates in cell migration and adhesion, and in platelet aggregati  
 C:Superfamily: thrombospondin 1; EGF homology; thrombospondin type 1 repeat homology;







GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 16, 2001, 16:13:12 ; Search time 11.88 Seconds  
(without alignments)  
910.650 Million cell updates/sec

Title: US-09-325-019-2

Sequence: 1 DFTNPAPLEDTSSRPQFCMKP.....NPNDIFADLESYPPFSEIAN 335

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 88757 seqs, 32294092 residues  
Total number of hits satisfying chosen parameters: 88757

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_39:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	869.5	44.9	349	1 CTGF_HUMAN	P29279 homo sapien
2	839.5	43.3	349	1 CTGF_PIG	O19113 homo scrofa
3	831.5	42.9	348	1 CTGF_MOUSE	P29266 mus musculus
4	793.5	41.0	349	1 CTGF_BOVIN	O18739 bos taurus
5	772.5	39.9	379	1 CYR6_MOUSE	P18406 mus musculus
6	764.5	39.5	381	1 CYR6_HUMAN	O00622 homo sapien
7	758	39.1	375	1 CE10_CHICK	P19336 gallus gall
8	743.5	38.4	351	1 NOV_CHICK	P28686 gallus gall
9	741.5	38.3	353	1 NOV_CHICK	P42642 coturnix co
10	726.5	37.5	354	1 NOV_MOUSE	O64299 mus musculus
11	726.5	37.5	357	1 NOV_HUMAN	P48745 homo sapien
12	703	36.3	343	1 NOV_XENLA	P51609 xenopus lae
13	151.5	7.8	5179	1 MUC2_HUMAN	O02817 homo sapien
14	150	7.7	2531	1 NTCL_MOUSE	O01705 mus musculus
15	150	7.7	2813	1 VWF_CANFA	O28295 canis fami
16	150	7.7	2813	1 VWF_HUMAN	P04275 homo sapien
17	147.5	7.6	1178	1 TSPI_CHICK	P35440 gallus gall
18	144.5	7.5	837	1 MUC1_RAT	P98089 rattus norv
19	144	7.4	1173	1 TSPI_XENLA	P35448 xenopus lae
20	142.5	7.4	1170	1 TSPI_HUMAN	P07996 homo sapien
21	140.5	7.3	1170	1 TSPI_MOUSE	P35441 mus musculus
22	139	7.2	1700	1 BAR3_CHITE	O03376 chironomus
23	138.5	7.2	2139	1 CRB_DROME	P10040 drosophila
24	138	7.1	555	1 DP87_DICDI	O04503 dicystostell
25	138	7.1	1056	1 MUC5_HUMAN	P98088 homo sapien
26	138	7.1	4289	1 TENX_HUMAN	P22105 homo sapien
27	137.5	7.1	810	1 NEIL_RAT	O62919 rattus norv
28	136	7.0	2437	1 NOTC_BRARE	P46530 brachydanio
29	135	7.0	2531	1 NTCL_RAT	O07008 rattus norv
30	134.5	6.9	854	1 LDIR_CRIGR	P35950 cricetus
31	134.5	6.9	864	1 LDIR_MOUSE	P35951 mus musculus
32	133.5	6.9	2444	1 NTCL_HUMAN	P46531 homo sapien
33	133	6.9	2524	1 NOTC_XENLA	P21783 xenopus lae

34	132	6.8	588	1 GRN_RAT	P23785 r granulin
35	131.5	6.8	1964	1 NTCL_MOUSE	P31695 mus musculus
36	130.5	6.7	1170	1 TSPI_BOVIN	O95116 bos taurus
37	129.5	6.7	1408	1 SERP_DROME	P18168 drosophila
38	129	6.7	1172	1 TSPI_HUMAN	P35442 homo sapien
39	128.5	6.6	2911	1 FBN2_HUMAN	P35556 homo sapien
40	128.5	6.6	4544	1 LRP1_HUMAN	O07954 homo sapien
41	128	6.6	3075	1 LRP1_MOUSE	P25391 homo sapien
42	126.5	6.5	1895	1 YLK3_DROME	P41951 caenorhabd
43	126.5	6.5	2703	1 NOTC_DROME	P07207 drosophila
44	126	6.5	810	1 NOTC_HUMAN	O92832 homo sapien
45	126	6.5	1150	1 APWU_PIG	P12021 sus scrofa

## ALIGNMENTS

RESULT	ID	CTGF_HUMAN	STANDARD:	PRT:	349 AA.
AC	P29279;				
DT	01-DEC-1992 (Rel. 24, Created)				
DT	01-DEC-1992 (Rel. 24, Last sequence update)				
DT	30-MAY-2000 (Rel. 39, Last annotation update)				
DE	CONNECTIVE TISSUE GROWTH FACTOR PRECURSOR.				
GN	CTGF.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE-UMBILICAL VEIN ENDOTHELIAL CELLS;				
RX	MEDLINE-91373462; PubMed-1654338;				
RA	Bradham D.M., Igarashi A., Potter R.L., Grotendorst G.R.;				
RT	"Connective tissue growth factor: a cysteine-rich mitogen secreted by				
RT	human vascular endothelial cells is related to the SRC-induced				
RL	immediate early gene product CPE-10.";				
RL	J. Cell Biol. 114:1285-1294(1991).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE-UMBILICAL VEIN ENDOTHELIAL CELLS;				
RX	MEDLINE-93187114; PubMed-1293144;				
RA	Igarashi A., Bradham D.M., Okochi H., Grotendorst G.R.;				
RT	"Connective tissue growth factor.";				
RL	J. Dermatol. 19:642-643(1992).				
RN	[3]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE-97207446; PubMed-9054739;				
RA	Demar B.S., Werner A., Gartner J.M., Do D.D., Godoy N., Nauck M.,				
RT	Marz W., Rupp J., Pech M., Luescher T.F.;				
RT	"Human connective tissue growth factor is expressed in advanced				
RL	atherosclerotic lesions.";				
CC	Circulation 95:831-839(1997).				
CC	- FUNCTION: MAJOR CONNECTIVE TISSUE MITOCHONDRIAL SECRETED BY				
CC	HUMAN VASCULAR ENDOTHELIAL CELLS. THIS IMMEDIATE-EARLY PROTEIN				
CC	MAY BIND ONE OF THE PDGF CELL SURFACE RECEPTORS.				
CC	- SUBUNIT: MONOMER.				
CC	- ALTERNATIVE PRODUCTS: A SHORTER FORM MAY BE PRODUCED BY				
CC	ALTERNATIVE SPLICING OF THE SAME GENE.				
CC	- SIMILARITY: BELONGS TO THE INSULIN-LIKE GROWTH FACTOR BINDING				
CC	PROTEIN FAMILY. CPE-10/CYR61/CTGF/ETSP-12/NOV PROTEIN SUBFAMILY.				
CC	- SIMILARITY: CONTAINS 1 VWF DOMAIN.				
CC	- SIMILARITY: CONTAINS 1 C-TERMINAL CYSTEINE KNOT-LIKE DOMAIN (CTCK).				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>				
CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).				
CC	EMBL; M92934; AAA91279.1; -				

DR EMBL: X78947; CA55544.1; -.  
DR PIR: A40551; A40551.  
DR PIR: S44205; S44205.  
DR MIM: 121009; -.  
DR INTERPRO: IPR000359; -.  
DR INTERPRO: IPR000867; -.  
DR INTERPRO: IPR000864; -.  
DR INTERPRO: IPR001007; -.  
DR PFAM: PF00007; Cys\_knot; 1.  
DR PFAM: PF00219; IGFBP; 1.  
DR PFAM: PF00090; tsp\_1; 1.  
DR PFAM: PF00093; wvc; 1.  
DR PROSITE: PS00222; IGF\_BINDING; 1.  
DR PROSITE: PS01185; CTCK\_1; 1.  
DR PROSITE: PS01225; CTCK\_2; 1.  
DR PROSITE: PS01208; WFC; 1.  
KW Growth factor binding; Signal; Alternative splicing.  
FT SIGNAL 1 26  
FT CHAIN 27 349  
FT DOMAIN 101 167  
FT DOMAIN 256 330  
FT DISULFID 256 293  
FT DISULFID 273 307  
FT DISULFID 284 323  
FT DISULFID 287 325  
FT DISULFID 292 329  
FT CARBOHD 28 28  
FT CARBOHD 225 225  
FT VARSPIC 172 198  
SQ SEQUENCE 349 AA; 38069 MW; 0BCEf8470B357EA95 CRC64;

Query Match 44.9%; Score 869.5; DB 1; Length 349;  
Best Local Similarity 45.6%; Pred. No. 1.2e-62;  
Matches 156; Conservative 53; Mismatches 98; Indels 35; Gaps 8;

OY 12 SRP---OPCKMPCECPSPR-PCPLGSLITDGCRCCKMCAQQLDNCNTEAICDPHGL 67  
DB 21 SRPAGQSCGSCGCPDPAKCPAGVSLVLDGCGCCRCVCAQQLDELCTERDPCDPHGL 80  
OY 68 YCDYSGDRPRAYIGCAOVGVGLDGVRYNNGSFQPNCKYNTCTIDGAVGCTPLC-L 126  
DB 81 FCDY-GSPANKRIGVCTAKDGAFCVFGTVYRSBSFSQSCYQCTCLDGAAGCVPILSM 139  
OY 127 RYRPRRLMCPHRRVSTIGHCCCEQWCEDDAKRPRTAPRTDGAFAVGEVAMR---- 182  
DB 140 DYRLSPDPCPPRRYKLTGKCCEEWVDE-----PKDQ---TVVGPAALAAVRLDRT 187  
OY 183 -----NCIATSPMSPCSTSGLGVSRTISNVNAOCMPQESRLCNLRPCDVDIH 232  
DB 188 RGPDPMTMRANCLVOTTEMASCKTCGMGISTRTVNDNAFCRLKQSRICMVRPCEADLE 247  
OY 233 TLIRAGKICLAVOPEASNFLLACISTRSYOPKCYGCVMQDNRCICITYKSTIDVSFOC 292  
DB 248 ENIKKGGKICIRPKISKIRKIFELSCSTSMKITYRAKFCVCTDGRCTHRTITLVEVERK 307  
OY 293 PDGIGFSRQVIMINACFCNLSCRPNDIFADL---ESYDPDS 331  
DB 308 PDGEVMAKKMMPIKTCACHYNCPPGNDIFESLYYRKMTGDM 349  
RESULT 2  
CTGF\_PIG STANDARD: PRT: 349 AA.  
AC 019113;  
DT 15-JUL-1998 (Rel. 36, Created)  
DT 15-JUL-1998 (Rel. 36, Last sequence update)  
DT 15-JUL-1998 (Rel. 36, Last annotation update)  
DE CONNECTIVE TISSUE GROWTH FACTOR PRECURSOR.  
GN CTGF.  
OS Sus scrofa (pig).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.

RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE-UTERUS;  
RX MEDLINE=97390475; PubMed=9242708;  
RA Brigstock D.R., Steffen C.L., Kim G.Y., Vegunta R.K., Diehl J.R.,  
RA Harding P.A.;  
RT "Purification and characterization of novel heparin-binding growth  
RT factors in uterine secretory fluids. Identification as heparin-  
RT regulated Mr 10,000 forms of connective tissue growth factor.";  
RL J. Biol. Chem. 272:20275-20282(1997).  
CC -1- FUNCTION: MAJOR CONNECTIVE TISSUE MITOATTRACTANT SECRETED BY  
CC HUMAN VASCULAR ENDOTHELIAL CELLS. THIS IMMEDIATE-EARLY PROTEIN  
CC MAY BIND ONE OF THE PDGF CELL SURFACE RECEPTORS.  
CC -1- SUBUNIT: MONOMER (BY SIMILARITY).  
CC -1- SIMILARITY: BELONGS TO THE INSULIN-LIKE GROWTH FACTOR BINDING  
CC PROTEIN FAMILY. CER-10/CYR61/CTRG/ETSP-12/NOV PROTEIN SUBFAMILY.  
CC -1- SIMILARITY: CONTAINS 1 WFC DOMAIN.  
CC -1- SIMILARITY: CONTAINS 1 C-TERMINAL CYSTINE KNOT-LIKE DOMAIN (CTCK).  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.ebi.ac.uk/announcements>  
CC or send an email to [license@ebi.ac.uk](mailto:license@ebi.ac.uk)).  
CC -----

DR EMBL: U83916; AAC48756.1; -.  
DR INTERPRO: IPR000359; -.  
DR INTERPRO: IPR000864; -.  
DR INTERPRO: IPR000867; -.  
DR INTERPRO: IPR001007; -.  
DR PFAM: PF00007; Cys\_knot; 1.  
DR PFAM: PF00219; IGFBP; 1.  
DR PFAM: PF00090; tsp\_1; 1.  
DR PFAM: PF00093; wvc; 1.  
DR PROSITE: PS00222; IGF\_BINDING; 1.  
DR PROSITE: PS01185; CTCK\_1; 1.  
DR PROSITE: PS01225; CTCK\_2; 1.  
DR PROSITE: PS01208; WFC; 1.  
KW Growth factor binding; Signal.  
FT SIGNAL 1 26  
FT CHAIN 27 349  
FT DOMAIN 101 167  
FT DOMAIN 256 330  
FT DISULFID 256 293  
FT DISULFID 273 307  
FT DISULFID 284 323  
FT DISULFID 287 325  
FT DISULFID 292 329  
SQ SEQUENCE 349 AA; 38007 MW; BB510E2B2B52D4A0 CRC64;

Query Match 43.3%; Score 839.5; DB 1; Length 349;  
Best Local Similarity 44.7%; Pred. No. 3.1e-60;  
Matches 153; Conservative 53; Mismatches 101; Indels 35; Gaps 8;

OY 12 SRP---OPCKMPCECPSPR-CPPLGSLITDGCRCCKMCAQQLDNCNTEAICDPHGL 67  
DB 21 SRPAGQSCGSCGCAAKRRACRPAGVSLVLDGCGCCRLCAKQJLDELCTERDPCDPHGL 80  
OY 68 YCDYSGDRPRAYIGCAOVGVGLDGVRYNNGSFQPNCKYNTCTIDGAVGCTPLC-L 126  
DB 81 FCDY-GSPANKRIGVCTAKDGAFCVFGTVYRSBSFSQSCYQCTCLDGAAGCVPILSM 139  
OY 127 RYRPRRLMCPHRRVSTIGHCCCEQWCEDDAKRPRTAPRTDGAFAVGEVAMR---- 182  
DB 140 DYRLSPDPCPPRRYKLTGKCCEEWVDE-----PKD---HTVVGPAALAAVRLDRT 187  
OY 183 -----NCIATSPMSPCSTSGLGVSRTISNVNAOCMPQESRLCNLRPCDVDIH 232  
DB 188 RGPDPMTMRANCLVOTTEMASCKTCGMGISTRTVNDNAFCRLKQSRICMVRPCEADLE 247

```

FT DISULFID 255 292 BY SIMILARITY.
FT DISULFID 272 306 BY SIMILARITY.
FT DISULFID 283 322 BY SIMILARITY.
FT DISULFID 286 324 BY SIMILARITY.
FT DISULFID 291 328 BY SIMILARITY.
FT CONFLICT 161 161 K -> E (IN REF. 2).
SO SEQUENCE 348 AA; 37793 MW; 735B65B6A711686F CMC64;

Query Match 42.9%; Score 831.5; DB 1; Length 348;
Best Local Similarity 44.2%; Pred. No. 1,3e-59;
Matches 151; Conservative 56; Mismatches 100; Indels 35; Gaps 8;

OY 12 SRP---QFCMKPCBC-PSPSPRCPLGVSLITDGCCECKMCAQOLGDNCTEAICDPHHGL 67
   || | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 20 TRPATGDCCSAQCOCFAEAPHCPSAGVSLVLVDGCCGCCCHVCAKQLDELCTERPPCDPHKGL 79

OY YCDSGDRPRALIGCAQAVGVGCLDGVRVNNNGSPDPNCKYNTCTIDGAVGCPRLC-L 126
   || | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 80 FCDPE-GSPANRKRIGCTAKDKAGPCVFEGSGYSRGSEFSOSCKTQCTCLDGAAGCVPLCSM 138

OY RVRRPRLCPHRRRSISGHCCSQWVCEDDAKRPRTKPTAPRDTGAFDAVEAEAMHR---- 182
   || | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 139 DVRLRSPDCPFPRRKRLGKCCKEKWVCE-----PKDR---TVNPGPALAAVRLDET 186

OY 183 -----NCIAVTSPWSPCSSTSGLGAVSTRISNVNAQCMPDESRLCNLRPCDVDIH 232
   || | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 187 FGPDPTMRANCIQVTETMSACSKTCGMGISIRVNDMTFCLEKXOSRLCMWRPCEADLE 246

OY 233 TLTRAGKKCIADVOPESAAMNFLTAGCISTRSYQRPYCGVCMNRCCIYYKSKTIDVSFOC 292
   || | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 247 ENIKGGKKCFIETPKLAKPVKEFLSGCTSVKTYRAKFCEVCTGDGRCPHTRTTLTPVEREK 306

OY 293 PDGLGFSSROVLIMINACFNLSGRANDNFADL---ESYPDEF 331
   || | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 307 PDGEIMKMMFMFIKTACACHYNCPCGNDNDFESLYRNKMIGDMA 348

RESULT 4
CTGF_BOVIN STANDARD: PRI: 349 AA.
ID CTGF_BOVIN
AC O18739;
AC 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE CONNECTIVE TISSUE GROWTH FACTOR PRECURSOR.
GN CTGF.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
RN [1]
RP SEQUENCE FROM N.A.
RA TISSUE-AORTA;
RA Lillienstiek B., Lin Z., Fotsis T., Schlimski M., Bierhaus A.,
RA Kanitz M., Kaufmann G., Schweigert L., Ziegler R., Nawroth P.P.;
RL Submitted (Aug-1997) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: MAJOR CONNECTIVE TISSUE MITOGENATTRACTANT SECRETED BY
CC HUMAN VASCULAR ENDOTHELIAL CELLS. THIS IMMEDIATE-EARLY PROTEIN
CC MAY BIND ONE OF THE PGF CELL SURFACE RECEPTORS.
CC -!- SUBUNIT: MONOMER (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE INSULIN-LIKE GROWTH FACTOR BINDING
CC PROTEIN FAMILY. CER-10/CYR61/CTGF/FISP-12/NOV PROTEIN SUBFAMILY.
CC -!- SIMILARITY: CONTAINS 1 VMBC DOMAIN
CC -!- SIMILARITY: CONTAINS 1 C-TERMINAL CYSTINE KNOT-LIKE DOMAIN (CTCK).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```

```

CC -----
DR EMBL: AF000137; AAB6596.1; -
DR INTERPRO: IPR000359; -
DR INTERPRO: IPR000867; -
DR INTERPRO: IPR000884; -
DR INTERPRO: IPR001007; -
DR PFAM: PF000007; Cys_knot; 1.
DR PFAM: PF00219; IGFBP; 1.
DR PFAM: PF00093; tsp_1; 1.
DR PFAM: PF00090; tsp_1; 1.
DR PROSITE: PS00222; IGF_BINDING; FALSE_NEG.
DR PROSITE: PS01185; CTCK_1; 1.
DR PROSITE: PS01225; CTCK_2; 1.
DR PROSITE: PS01208; WVEC; FALSE_NEG.
DR PROSITE: PS01208; WVEC; FALSE_NEG.
KW Growth factor binding; Signal.
FT CHAIN 1 26 POTENTIAL.
FT SIGNAL 27 349 CONNECTIVE TISSUE GROWTH FACTOR.
FT DOMAIN 101 167 WVEC.
FT DOMAIN 256 330 CTCK.
FT DISULFID 256 293 BY SIMILARITY.
FT DISULFID 273 307 BY SIMILARITY.
FT DISULFID 287 325 BY SIMILARITY.
FT DISULFID 292 329 BY SIMILARITY.
SO SEQUENCE 349 AA; 38152 MM; D919023AE40D212E CRC64;

```

```

Query Match 41.0%; Score 793.5; DB 1; Length 349;
Best Local Similarity 43.1%; Pred. No. 1.4e-56;
Matches 148; Conservative 52; Mismatches 106; Indels 37; Gaps 8;

```

```

OY 12 SRP-----QFCWPCPCPPSP-PRCLGLVSLITDGECCCKMCAQOLGDMCTEALCDPRG 66
DB 21 SRPASSQDCSPCCPGAPCPAGVSLVLDGGCC-VCAKOLSELECTERDPCDPRKG 79
OY 67 LYCDYSGDRPRAISGCAQVGVGLDGVRYNNGOSFQPMCKNYCTCIDGAVGCTPRC- 125
DB 80 LFCDF-GSPTRNKIGVCAKADGAPYIPFGTYQSESFQSSCKRYCTCIDGSGVGLPCL 138
OY 126 LVRPRLMCPHRRVSIIPGHCCFQWCCEDAKRPRKTAAPDGTGAFAVEVEAMHR--- 182
DB 139 VDVRLPSPDCFPFRVHKLPKCCCEMVSDEKEH-----TYVGPALAAVRIED 186
OY 183 -----NCIAVTSPMSPCSTSGGLGVSTRISNNAOCMPROESRLCLARCVDVI 231
DB 187 TFGPPTMIRACQVOTEMSAVSTKCMGSTRVTNNACRLEKQSLCVRCEADL 246
OY 232 HTLAKGRKCLAVQPEASMTTLAIGISTRSYOPKYCGVCAQMDNRCIPYKSKTIDVFSQ 291
DB 247 ENNIKKGKCIKTRPKISPIKFOISGCTSMKTYRAKFFGVCTDGCCTPRHTTILPVERK 306
OY 292 CPDGLGFSROYLWLNACFCNTLSCRNPNDIFADL--ESYPDFS 331
DB 307 CPDGEVMMKSMHFIKTCACHYNCPCDNDIFSLYRKMYGDMA 349

```

```

RESULT 5
CYR6_MOUSE STANDARD; PRT; 379 AA.
AC P18406;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE CYR61 PROTEIN PRECURSOR (3chr61).
GN CYR61 OR IGFBP10.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BAIB/C; TISSUE=FIBROBLAST;
RX MEDLINE=90287146; PubMed=2355916;
RA O'Brien T.P., Yang G.P., Sanders L., Lau L.F.;
RT "Expression of cyr61, a growth factor-inducible immediate-early

```

```

RT gene.";
RN Mol. Cell. Biol. 10:3569-3577(1990).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=BAIB; TISSUE=EMBRYONIC FIBROBLAST;
RX MEDLINE=91288203; PubMed=2062642;
RA Latinkic B.V., O'Brien T.P., Lau L.F.;
RT "Promoter function and structure of the growth factor-inducible
RT immediate early gene cyr61.";
RL Nucleic Acids Res. 19:3261-3267(1991).
CC -1- FUNCTION: MAY ACT AS ONE OF THE MANY GROWTH FACTOR-BINDING
CC PROTEINS. PROMOTES PROLIFERATION, MIGRATION AND ADHESION.
CC -1- TISSUE SPECIFICITY: LOW IN KIDNEY, ADRENAL GLAND, TESTES, BRAIN,
CC AND OVARY, MODERATE IN HEART, UTERUS, AND SKELETAL MUSCLE, HIGHEST
CC IN LUNG.
CC -1- DEVELOPMENTAL STAGE: EXPRESSED FROM G(0)/G(1) THROUGH MID-G(1) IN
CC NORMAL CELLS, AND AT A CONSTANT LEVEL IN RAPIDLY GROWING CELLS.
CC -1- INDUCTION: BY GROWTH FACTORS.
CC -1- SIMILARITY: BELONGS TO THE INSULIN-LIKE GROWTH FACTOR BINDING
CC PROTEIN FAMILY. CEF-10/CYR61/CTFG/FTSP-12/NOV PROTEIN SUBFAMILY.
CC -1- SIMILARITY: CONTAINS 1 C-TERMINAL CYSTEINE KNOT-LIKE DOMAIN (CTCK).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.1ab-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----

```

```

DR EMBL: M32490; AAA37512.1; -
DR EMBL: X56790; CAA40109.1; -
DR PIR: A35669; A35669.
DR MGD: MGI:88613; CYR61.
DR INTERPRO: IPR000359; -
DR INTERPRO: IPR000867; -
DR INTERPRO: IPR000884; -
DR INTERPRO: IPR001007; -
DR PFAM: PF000007; Cys_knot; 1.
DR PFAM: PF00219; IGFBP; 1.
DR PFAM: PF00093; tsp_1; 1.
DR PFAM: PF00090; tsp_1; 1.
DR PROSITE: PS00222; IGF_BINDING; 1.
DR PROSITE: PS01185; CTCK_1; 1.
DR PROSITE: PS01225; CTCK_2; 1.
DR PROSITE: PS01208; WVEC; 1.
KW Growth factor binding; Signal.
FT SIGNAL 1 24 POTENTIAL.
FT CHAIN 25 379 CYR61 PROTEIN.
FT DOMAIN 98 164 WVEC.
FT DOMAIN 284 358 CTCK.
FT DISULFID 284 321 BY SIMILARITY.
FT DISULFID 301 335 BY SIMILARITY.
FT DISULFID 312 351 BY SIMILARITY.
FT DISULFID 315 353 BY SIMILARITY.
FT DISULFID 320 357 BY SIMILARITY.
SO SEQUENCE 379 AA; 41709 MM; FA6B501AB56A8EE9 CRC64;

```

```

Query Match 39.9%; Score 772.5; DB 1; Length 379;
Best Local Similarity 39.6%; Pred. No. 7.3e-55;
Matches 141; Conservative 56; Mismatches 114; Indels 45; Gaps 6;

```

```

OY 17 CKNWPCPCPPSPPRCLGLVSLITDGECCCKMCAQOLGDMCTEALCDPRGLYCDVSGSRP 76
DB 26 CPAACHCPLKAPKCAPGVGLVNDGGCCVCAKOLNEDCSKTQPCDHTKGLCNP-GASS 84
OY 77 RYALGVC-AQVAVGCVLDGVRYNNGOSFQPMCKNYCTCIDGAVGCTPRC-LRVPRPLM 134
DB 85 TALNGICRAQSGRCFENSHYNGSEFQPMCKNYCTCIDGAVGCTPRCQGLSLPRLG 144
OY 135 CPHRRVSIIPGHCCFQWCCEDAKRPRKTAAPDGTGAFAVEVEAMHRN----- 183

```





SQ SEQUENCE 351 AA; 38268 MW; 1ECB3FA30586797 CRC64;

Query Match 38.4%; Score 743.5; DB 1; Length 351;

Best Local Similarity 41.2%; Pred. No. 1.4e-52;

Matches 135; Conservative 49; Mismatches 117; Indels 27; Gaps 7;

QY 9 DTSRRPQFCRMPK--ECPPSPRCPLGVSLLITDGCCECKMKCAQQLGDNCTEAALICDPHRG 66  
 DB 23 EVSGREACPRPGRCRPAEPAPCAVPVAVLDGCGCCCLVCARQRESCSPLLPDSEGG 82  
 QY 67 LYCDYSGDRPRRAIGVCAOVYGVGVLDGVRYNNGSFQFNCKYKNTCTIDGAVGCTPLC- 125  
 DB 83 LYCD-RGPEDEGGAGICVMLEBNCVFDGMYNNGETFPQSCYQCTCRGQIGCLPRCN 141  
 QY 126 LRYRPRLMCPHPRRVSTFGHCCQWVCEDDAKRPRTAPRD--TGADAVG----- 175  
 DB 142 LGILLPGDCCPPPRKLEVEGCECKWVCD-----PRDEVLLGGFAMAARYOEAATL 191  
 QY 176 --EVEAMHNCIAVTSWSPSCSTGCGVSTRISNVNAQCMPQESRLCNLRPCVDYDHT 233  
 DB 192 GIDVSSANCIQETTEMWASCSKSCGMGFSTRVTRNQCEWYKQTRLCMMRPECNE-EP 250  
 QY 234 LTKAGKCLAVYQPEASMTFLAGCISTRSYOPRYGVCVMDNRCCLIPYKSKTIDVSEFQCP 293  
 DB 251 SDRKGGKCIQTKSMKAVREYKNCISVQYTKRPYCGICLNDGRCTCPHNTKIQVEFRCP 310  
 QY 294 DGLGFSROYLWIMNACFCNLSCRPNDF 321  
 DB 311 QGKFLKPMMLINTVCVCHGNCPSNNAF 338

RESULT 9

NOV\_COTJA STANDARD; PRT; 353 AA.  
 AC P42642;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 01-OCT-1996 (Rel. 34, Last annotation update)  
 DE NOV PROTEIN PRECURSOR.  
 GN Colunrix cotunrix japonica (Japanese quail).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
 OC Colunrix.  
 RN 11  
 RP SEQUENCE FROM N.A.  
 RA Weiskirchen R., Bister K.;  
 RL Submitted (Aug-1994) to the EMBL/Genbank/DBJ databases.  
 CC -1- FUNCTION: IMMEDIATE-EARLY PROTEIN LIKELY TO PLAY A ROLE IN CELL  
 CC GROWTH REGULATION (BY SIMILARITY).  
 CC -1- SIMILARITY: BELONGS TO THE INSULIN-LIKE GROWTH FACTOR BINDING  
 CC PROTEIN FAMILY. CEF-10/CYR61/CTFG/FISP-12/NOV PROTEIN SUBFAMILY.  
 CC -1- SIMILARITY: CONTAINS 1 VMFC DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 1 C-TERMINAL CYSTINE KNOT-LIKE DOMAIN (CTCK).  
 CC OS This Swiss-Prot entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL: U13063; AAA21128.1; -  
 DR INTERPRO: IPR000359; -  
 DR INTERPRO: IPR000867; -  
 DR INTERPRO: IPR000884; -  
 DR INTERPRO: IPR001007; -  
 DR PFAM: PF00007; Cys\_knot; 1.  
 DR PFAM: PF00219; IGFBP; 1.  
 DR PFAM: PF00090; tsp\_1; 1.  
 DR PFAM: PF00093; vwc; 1.

DR PROSITE: PS00222; IGF BINDING; 1.  
 DR PROSITE: PS01185; CTCK\_1; 1.  
 DR PROSITE: PS01225; CTCK\_2; 1.  
 DR PROSITE: PS01208; VMFC; 1.  
 KW Proto-oncogene; Growth factor binding; Signal.  
 FT SIGNAL 1 26  
 FT CHAIN 27 353  
 FT DOMAIN 106 172  
 FT DOMAIN 260 334  
 FT DOMAIN 260 334  
 FT DISULFID 260 334  
 FT DISULFID 277 311  
 FT DISULFID 288 327  
 FT DISULFID 291 329  
 FT DISULFID 296 333  
 FT CARBOHYD 276 276  
 SQ SEQUENCE 353 AA; 38667 MW; 711D9F8533882E89 CRC64;

Query Match 38.3%; Score 741.5; DB 1; Length 353;  
 Best Local Similarity 41.2%; Pred. No. 2e-52;  
 Matches 135; Conservative 49; Mismatches 117; Indels 27; Gaps 7;

QY 9 DTSRRPQFCRMPK--ECPPSPRCPLGVSLLITDGCCECKMKCAQQLGDNCTEAALICDPHRG 66  
 DB 25 EVNGREAPCRPGRCRPAEPAPCAVPVAVLDGCGCCCLVCARQRESCSPLLPDSEGG 84  
 QY 67 LYCDYSGDRPRRAIGVCAOVYGVGVLDGVRYNNGSFQFNCKYKNTCTIDGAVGCTPLC- 125  
 DB 85 LYCD-RGPEDEGGTIGICVMLEBNCVFDGMYNNGETFPQSCYQCTCRGQIGCLPRCN 143  
 QY 126 LRYRPRLMCPHPRRVSTFGHCCQWVCEDDAKRPRTAPRD--TGADAVG----- 175  
 DB 144 LGILLPGDCCPPPRKLEVEGCECKWVCD-----PRDEVLLGGFAMAARYOEAATL 193  
 QY 176 --EVEAMHNCIAVTSWSPSCSTGCGVSTRISNVNAQCMPQESRLCNLRPCVDYDHT 233  
 DB 194 GIDVSSANCIQETTEMWASCSKSCGMGFSTRVTRNQCEWYKQTRLCMMRPECNE-EP 252  
 QY 234 LTKAGKCLAVYQPEASMTFLAGCISTRSYOPRYGVCVMDNRCCLIPYKSKTIDVSEFQCP 293  
 DB 253 SDRKGGKCIQTKSMKAVREYKNCISVQYTKRPYCGICLNDGRCTCPHNTKIQVEFRCP 312  
 QY 294 DGLGFSROYLWIMNACFCNLSCRPNDF 321  
 DB 313 QGKFLKPMMLINTVCVCHGNCPSNNAF 340

RESULT 10

NOV\_MOUSE STANDARD; PRT; 354 AA.  
 AC Q64299;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE NOV PROTEIN HOMOLOG PRECURSOR (NOVH).  
 GN Mus.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 RN 11  
 RP SEQUENCE FROM N.A.  
 RA STRAIN-129/SV, AND ICR; TISSUE-BRAIN;  
 RC MEDLINE-97131523; PubMed-8975721;  
 RX Snailh M.R., Natarajan D., Taylor L.B., Choi C.P., Martinerie C.,  
 RT Perbal B., Schofield P.N., Boulter C.A.;  
 RL "Genomic structure and chromosomal mapping of the mouse nov gene";  
 RL Genomics 38:425-428(1996).  
 RN  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-C57BL/6;  
 RX MEDLINE-96204003; PubMed-8622864;  
 RA Martinerie C., Chevalier G., Rauscher F.J. III, Perbal B.;  
 RT "Regulation of nov by Wt1: a potential role for nov in





Query Match	37.5%	Score 726.5	DB 1	Length 357
Best Local Similarity	42.0%	Pred. No. 3.3e-51		
Matches 137	Conservative 46	Mismatches 116	Indels 27	Gaps 8
Qy	10	TSRPFQCKMPCDEPPSPRCPLGSLTLTDCGCECCMKMAQOLGDNCTAATCDPHRGHYC	69	
Db	32	TQRCPPQC - GRCPCATPTPCAPYRAVLDDGSCCLTCVARGGESCSDLPEDDESSGLYC	89	
Qy	70	DYSGDRPYRATGCAQVAVGVGCVLDGVRYNMGSGSPQPKNCTCTDCAVGCTPLC-LRV	128	
Db	90	DRSAD-PSNONGICPAVAGDNCVEGIVYRSGEKFPSCRCQCTCRDQIGCVPRCOLDY	148	
Qy	129	RPPRLMCPPHPRVSIIPGHCCEDQWVC---EDD-----AKRPRKTAAPDRTGAFDVG-E	176	
Db	149	LLPEPNCAPKPKVEYEPGECEKVIICGPBEDBSLGLTLTAIRPEAT-----LQVE	198	
Qy	177	VEAHMRCIAVTSPPWSPCSTSGLGVSSTRISNVANAQCPMEQESRLCNLRPCDVI-IHTLI	235	
Db	199	VSDSVSNVICEDTETWTACSKSGMGFSRVRNRRNGCEMLKQTRLCVAPRCQDEQEPD	258	
Qy	236	KAKKCLAVYQPEASAMNTTLGCTISRYQRYGCVCDNRCCITPYKSKTIDVSQCDDG	295	
Db	259	KKKKCLCTTKSKLKHQFNKCTSLAHYKBPFCGVCSDRCQTPHNTKTIOAEFCSPG	318	
Qy	296	LGFSROVLMINACFCNLSCRPNDFE	321	
Db	319	QIVKKPVAIVICTCTCHTNCPRKNEAF	344	
RESULT 12				
NOV_XENLA				
ID	NOV_XENLA	STANDARD:	PRT:	343 AA.
AC	P51609:			
DT	01-OCT-1996 (Rel. 34, Created)			
DT	01-OCT-1996 (Rel. 34, Last sequence update)			
DT	01-OCT-1996 (Rel. 34, Last annotation update)			
DE	NOV PROTEIN HOMOLOG PRECURSOR (XNOV).			
GN	NOV.			
OS	Xenopus laevis (African clawed frog).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;			
OC	Xenopodinae; Xenopus.			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=96257272; PubMed=8666280;			
RA	Ying Z., King M.L.;			
RT	"Isolation and characterization of xnov, a Xenopus laevis ortholog of			
RL	the chicken nov gene.";			
RL	Gene 171:243-248(1996).			
CC	-1- FUNCTION: IMMEDIATE-EARLY PROTEIN LIKELY TO PLAY A ROLE IN CELL			
CC	GROWTH REGULATION (BY SIMILARITY).			
CC	-1- SIMILARITY: BELONGS TO THE INSULIN-LIKE GROWTH FACTOR BINDING			
CC	PROTEIN FAMILY. CER-10/CYR61/CTGF/FTSP-12/NOV PROTEIN SUBFAMILY.			
CC	-1- SIMILARITY: CONTAINS 1 WFEC DOMAIN.			
CC	-1- SIMILARITY: CONTAINS 1 C-TERMINAL CYSTINE KNOT-LIKE DOMAIN (CTCK).			
CC				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (see <a href="http://www.isb-sib.ch/announce/a">http://www.isb-sib.ch/announce/a</a>			
CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).			
CC				
DR	EMBL: U37063; AAB17096.1; -			
DR	INTERPRO: IPR000359; -			
DR	INTERPRO: IPR000867; -			
DR	INTERPRO: IPR000884; -			
DR	INTERPRO: IPR001007; -			
DR	PFAM: PF00007; Cys_knot; 1.			
DR	PFAM: PF00219; IGFBP; 1.			

Query Match	Best Local Similarity	Score 703;	DB 1;	Length 343;
Matches 133;	Conservative 50;	Mismatches 110;	Indels 26;	Gaps
15	QPCKMP-CBCEPSPRCLGSLITDGCBCCKMAQQLDNCETAICDPHAGLYCDYSG	73		
19	QKCPQCDQCEPEEPSCAPSVLLIIDLGGCCCPVCARQEGESHLNPPCEDKGLCFEFA	78		
74	DPRRAIGVCAVGVGCVLDGVRNNNGSPDPNCKYNTCTCIDGAVGCPPLC-LVVRPPR	132		
79	D-PRNETTCYCALGENSCVFDGVVVRNRRSEFQPSCKYHCTCLNGHIGVPCRNLDLLPG	137		
133	LMCPHRRYSIRGHCCCEQVEDD-----AKRPRTAPRPDGTGAFDAVGEVEMH	181		
138	PDCPPRRKRVYECCEKRVCDSKEMAIIGFAMAAYRPAAL-----GIDASDTSFA--	190		
162	RNCIAYTSPWSPCSTSGGLGVSTRISVNAQCPQESRILMLRPDQVDITLI-KAKRK	240		
191	--CIAQTTEWSACSTKCGKGVSSRVTNRNARCEMOQIILCVWRACEEPGWHVEKGRK	248		
241	CLAVQPEASNMFTLAGCSTFSTRYQPKVCGVCMQDRCCIPRYSKITIDVSFOCPDGLGFSR	300		
249	CVRRKTKTKRPHFHKNTCSVQPYRKFCGQSDGRCCTPHSTKTMHVEVCPQKRIYVK	308		
301	QVLTINACFCNLSCRPNPD 319			
309	PVMVISTCVCHYNC--POD 325			

RESULT 13

ID	MUC2_HUMAN	STANDARD:	PRT: 5179 AA.
AC	Q02817; Q14878;		
DT	01-JUN-1994 (Rel. 29, Created)		
DT	01-NOV-1997 (Rel. 35, Last sequence update)		
DT	15-JUL-1999 (Rel. 38, Last annotation update)		
DE	MUCIN 2 PRECURSOR (INTESTINAL MUCIN 2).		
GN	MUC2 OR SMUC.		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.		
RE	[1]		
RE	SEQUENCE FROM N.A.		
RC	TISUB-INTERESTINE;		
RX	MEDLINE=941132002; PubMed=8300571;		
RT	Gum J.R., Jr., Hicks J.W., Toribara N.W., Siddiki B., Kim Y.S.;		
RT	"Molecular cloning of human intestinal mucin (MUC2) cDNA.		
RT	Identification of the amino terminus and overall sequence similarity		
RT	to prepro-von Willebrand factor";		
RL	J. Biol. Chem. 269:2440-2446(1994).		
RL	[2]		
RN	SEQUENCE OF 626-1895 AND 4196-5179 FROM N.A.		
RC	TISSUE-COLON;		

```

RX MEDLINE=93016075; PubMed=1400449;
RA Gum J.R., Jr., Hicks J.W., Toribara N.W., Rothe E.-M., Lagace R.E.,
RA Kim Y.S.;
RT "The human MUC2 intestinal mucin has cysteine-rich subdomains located
RT both upstream and downstream of its central repetitive region.";
RL J. Biol. Chem. 267:21375-21383(1992).
RP [3]
RP SEQUENCE OF 1343-1895 AND 4176-4195 FROM N.A.
RA MEDLINE=91358717; PubMed=1885763;
RA Toribara N.W., Gum J.R., Jr., Culhane P.J., Lagace R.E., Hicks J.W.,
RA Petersen G.M., Kim Y.S.;
RT "MUC-2 human small intestinal mucin gene structure. Repeated arrays
RT and polymorphism.";
RL J. Clin. Invest. 88:1005-1013(1991).
CC -1- FUNCTION: COATS THE EPITHELIA OF THE INTESTINES, AIRWAYS, AND
CC OTHER MUCUS MEMBRANE-CONTAINING ORGANS. THOUGHT TO PROVIDE A
CC PROTECTIVE, LUBRICATING BARRIER AGAINST PARTICLES AND INFECTIONS
CC AGENTS AT MUCOSAL SURFACES.
CC -1- SUBUNIT: MULTIMERIC.
CC -1- SUBCELLULAR LOCATION: SECRETED.
CC -1- TISSUE SPECIFICITY: COLON, SMALL INTESTINE, COLONIC TUMORS,
CC BRONCHUS, CERVIX AND GALL BLADDER.
CC -1- PTM: ALL CYSTEINE RESIDUES ARE INVOLVED IN INTRACHAIN OR
CC INTERCHAIN DISULFIDE BONDS (BY SIMILARITY).
CC -1- POLYMORPHISM: THE NUMBER OF REPEATS IS HIGHLY POLYMORPHIC AND
CC VARIES AMONG DIFFERENT ALLELES.
CC -1- SIMILARITY: THE N-TERMINAL DOMAIN SHOWS SOME SIMILARITY TO THAT
CC OF SILKWORK HEMOCYTIN.
CC -1- SIMILARITY: CONTAINS 1 C-TERMINAL CYSTEINE KNOT-LIKE DOMAIN (CTCK).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; L2198; AAB95295.1; -
DR EMBL; M74027; AAB59875.1; -
DR EMBL; M94131; AAB59163.1; -
DR EMBL; M94132; AAB59164.1; -
DR MIM; 158370; -
DR INTERPRO: IPR000359; -
DR INTERPRO: IPR000561; -
DR INTERPRO: IPR001007; -
DR INTERPRO: IPR001846; -
DR INTERPRO: IPR002400; -
DR PFM; PFI00007; Cys_knot; 1.
DR PFM; PFI00094; vwd; 4.
DR PRINTS; PR00438; GFCYSKNOT.
DR PROSITE; PS00022; EGF_1; UNKNOWN_1.
DR PROSITE; PS01185; CTCK_1; 1.
DR PROSITE; PS01225; CTCK_2; 1.
DR PROSITE; PS01208; vWFC; 2.
KW Glycoprotein; Repeat; Polymorphism; Signal.
FT SIGNAL 1 20
FT CHAIN 21 5179
FT DOMAIN 1401 1747 MUCIN 2.
FT REPEAT 1401 1416 APPROXIMATE REPEATS.
FT REPEAT 1417 1432 1.
FT REPEAT 1433 1448 2.
FT REPEAT 1449 1464 3.
FT REPEAT 1465 1471 4.
FT REPEAT 1472 1478 5.
FT REPEAT 1479 1494 6.
FT REPEAT 1495 1517 7A.
FT REPEAT 1518 1533 7B.
FT REPEAT 1534 1556 8A.
FT REPEAT 1557 1572 8B.
FT REPEAT 1573 1596 9A.
FT REPEAT 1597 1612 9B.
FT REPEAT 10A.

```

```

FT REPEAT 1613 1635 10B.
FT REPEAT 1636 1651 11A.
FT REPEAT 1652 1675 11B.
FT REPEAT 1676 1683 12.
FT REPEAT 1684 1699 13.
FT REPEAT 1700 1715 14.
FT REPEAT 1716 1731 15.
FT REPEAT 1732 1747 16.
FT DOMAIN 4815 4886 vWFC 1.
FT DOMAIN 4924 4991 vWFC 2.
FT DOMAIN 5075 5160 CTCK.
FT DISULFID 5075 5122 BY SIMILARITY.
FT DISULFID 5089 5136 BY SIMILARITY.
FT DISULFID 5098 5152 BY SIMILARITY.
FT DISULFID 5102 5154 BY SIMILARITY.
FT CARBOHYD 163 163 ?
FT CARBOHYD 423 423 N-LINKED (GLCNAc...) (POTENTIAL).
FT CARBOHYD 670 670 N-LINKED (GLCNAc...) (POTENTIAL).
FT CARBOHYD 770 770 N-LINKED (GLCNAc...) (POTENTIAL).
FT CARBOHYD 894 894 N-LINKED (GLCNAc...) (POTENTIAL).
FT CARBOHYD 1139 1139 N-LINKED (GLCNAc...) (POTENTIAL).
FT CARBOHYD 1154 1154 N-LINKED (GLCNAc...) (POTENTIAL).
FT CARBOHYD 1215 1215 N-LINKED (GLCNAc...) (POTENTIAL).
FT CARBOHYD 1230 1230 N-LINKED (GLCNAc...) (POTENTIAL).
FT CARBOHYD 1246 1246 N-LINKED (GLCNAc...) (POTENTIAL).
FT CARBOHYD 1787 1787 N-LINKED (GLCNAc...) (POTENTIAL).
FT CARBOHYD 1820 1820 N-LINKED (GLCNAc...) (POTENTIAL).
FT CARBOHYD 4339 4339 N-LINKED (GLCNAc...) (POTENTIAL).
FT CARBOHYD 4351 4351 N-LINKED (GLCNAc...) (POTENTIAL).
FT CARBOHYD 4362 4362 N-LINKED (GLCNAc...) (POTENTIAL).
FT CARBOHYD 4373 4373 N-LINKED (GLCNAc...) (POTENTIAL).
FT CARBOHYD 4422 4422 N-LINKED (GLCNAc...) (POTENTIAL).
FT CARBOHYD 4438 4438 N-LINKED (GLCNAc...) (POTENTIAL).
FT CARBOHYD 4502 4502 N-LINKED (GLCNAc...) (POTENTIAL).
FT CARBOHYD 4616 4616 N-LINKED (GLCNAc...) (POTENTIAL).
FT CARBOHYD 4627 4627 N-LINKED (GLCNAc...) (POTENTIAL).
FT CARBOHYD 4752 4752 N-LINKED (GLCNAc...) (POTENTIAL).
FT CARBOHYD 4787 4787 N-LINKED (GLCNAc...) (POTENTIAL).
FT CARBOHYD 4881 4881 N-LINKED (GLCNAc...) (POTENTIAL).
FT CARBOHYD 4888 4888 N-LINKED (GLCNAc...) (POTENTIAL).
FT CARBOHYD 4955 4955 N-LINKED (GLCNAc...) (POTENTIAL).
FT CARBOHYD 4970 4970 N-LINKED (GLCNAc...) (POTENTIAL).
FT CARBOHYD 5019 5019 N-LINKED (GLCNAc...) (POTENTIAL).
FT CARBOHYD 5038 5038 N-LINKED (GLCNAc...) (POTENTIAL).
FT CARBOHYD 5069 5069 N-LINKED (GLCNAc...) (POTENTIAL).
FT CONFLICT 1351 1351 H -> L (IN REF. 3).
FT CONFLICT 1412 1412 T -> S (IN REF. 3).
FT CONFLICT 1449 1449 L -> P (IN REF. 3).
FT CONFLICT 1504 1504 M -> T (IN REF. 3).
FT CONFLICT 4192 4192 G -> S (IN REF. 2).
SQ SEQUENCE 5179 AA; 540295 MW; 85CD7571FB9A5663 CRC64;

Query Match 7.8%; Score 151.5; DB 1; Length 5179;
Best Local Similarity 20.5%; Pred. No. 0.0028;
Matches 99; Conservative 40; Mismatches 156; Indels 189; Gaps 26;

```

```

OY 149 EOMVCEDDAKPRKTAAPRDGA-----PDAYGEVAMHR---NCIAYTSPMSPCSTSC 198
DB 4916 PFWCESKGCVCVGNAYOGSPYSSKDCVCTDKVDNNTLLNLTACTH--VPCWTSC 4973
OY 199 GLGVSTRISNVNAQCWPEDESRLCNLRPCVDIHTLIKACK-----KCLAVYQ 246
DB 4974 SPGF--ELMAPEGCCKKQHTHCITIRPD-NQHVILKPDGFKSDPKNNCTFFSCVAKIHN 5030
OY 247 PEAISM-----NFTLAGCT-STRSTYOPKYC-----269
DB 5031 QLTISSVNTICPNFDASICIPGSIITFMPNCGCCTCTPRNETRVPCSTVPYTTVEVSAGCT 5090
OY 270 -----GVC-----MDN--RCCTIPYKSTIDVGFQCPDGLGFSRQVLMIN 306
DB 5091 KTYVLAMNCSGSGCTFVWYSAKAQAALDHSKSCCKEKTQSREVLVSCPNGSLTHYTHIE 5150
OY 307 ACFC 310
DB 5151 SCOC 5154

RESULT 14
NTCL.MOUSE STANDARD: PRT: 2531 AA.
ID NTCL.MOUSE
AC 001705;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE NEUROGENIC LOCUS NOTCH HOMOLOG PROTEIN 1 PRECURSOR (NOTCH PROTEIN).
GN NOTCH1 OR NOTCH.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-EMBRYO;
RX MEDLINE-93194170; PubMed-8449489;
RA Franco del Amo F., Gendron-Maguire M., Swiatek P.J., Jenkins N.A.,
RT Copeland N.G., Gridley T.;
RT "Cloning, analysis, and chromosomal localization of Notch-1, a mouse
RT homolog of Drosophila Notch.";
RL Genomics 15:259-264(1993).
RN [2]
RP SEQUENCE OF 1551-2170 FROM N.A.
RC TISSUE-EMBRYO;
RX MEDLINE-93048835; PubMed-1425352;
RA Franco del Amo F., Smith D.E., Swiatek P.J., Gendron-Maguire M.,
RT Greenspan R.J., McMahon A.P., Gridley T.;
RT "Expression pattern of Notch, a mouse homolog of Drosophila Notch,
RT suggests an important role in early postimplantation mouse
RT development.";
RL Development 115:737-744(1992).
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -1- DEVELOPMENTAL STAGE: EXPRESSED ALMOST UNIFORMLY IN EARLY EMBRYOS.
CC -1- SIMILARITY: CONTAINS 36 EGF-LIKE DOMAINS.
CC -1- SIMILARITY: CONTAINS 6 ANK REPEATS.
CC -1- SIMILARITY: HIGH, WITH OTHER NOTCH-TYPE PROTEINS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ebi.ac.uk/announcements/
CC or send an email to license@ebi.ac.uk).
CC -----
CC EMBL, 211866; CA77941.1; -.
CC HSSP; P00740; IIXA.
CC MGD; MGI:97363; NOTCH1.
CC INTERPRO; IPR000152; -.
CC INTERPRO; IPR000561; -.

```

```

DR INTERPRO; IPR000800; -.
DR INTERPRO; IPR001438; -.
DR INTERPRO; IPR001881; -.
DR INTERPRO; IPR002110; -.
DR PFAM; PF00008; EGF; 35.
DR PFAM; PF00023; ank; 6.
DR PFAM; PF00066; notch; 3.
DR PRINTS; PR00010; EGFBLD.
DR PROSITE; PS00088; ANK_REPEAT; 2.
DR PROSITE; PS0297; ANK_REPEAT_REGION; 1.
DR PROSITE; PS00010; ASX_HYDROXYL; 22.
DR PROSITE; PS00022; EGF_1; 34.
DR PROSITE; PS01186; EGF_2; 27.
DR PROSITE; PS01187; EGF_CA; 21.
KW Differentiation; Neurogenesis; Repeat; ANK repeat; EGF-like domain;
KW Transmembrane; Signal; Glycoprotein.
FT SIGNAL 1 18
FT CHAIN 19 2531
FT DOMAIN 19 1725
FT TRANSMEM 1726 1746
FT DOMAIN 1747 2531
FT DOMAIN 20 1426
FT DOMAIN 20 58
FT DOMAIN 59 99
FT DOMAIN 102 139
FT DOMAIN 140 176
FT DOMAIN 178 216
FT DOMAIN 218 255
FT DOMAIN 257 293
FT DOMAIN 295 333
FT DOMAIN 335 371
FT DOMAIN 372 410
FT DOMAIN 412 450
FT DOMAIN 452 488
FT DOMAIN 490 526
FT DOMAIN 528 564
FT DOMAIN 566 601
FT DOMAIN 603 639
FT DOMAIN 641 676
FT DOMAIN 678 714
FT DOMAIN 716 751
FT DOMAIN 753 789
FT DOMAIN 791 827
FT DOMAIN 829 867
FT DOMAIN 869 905
FT DOMAIN 907 943
FT DOMAIN 945 981
FT DOMAIN 983 1019
FT DOMAIN 1021 1057
FT DOMAIN 1059 1095
FT DOMAIN 1097 1143
FT DOMAIN 1145 1181
FT DOMAIN 1183 1219
FT DOMAIN 1221 1265
FT DOMAIN 1267 1305
FT DOMAIN 1307 1346
FT DOMAIN 1348 1384
FT DOMAIN 1387 1426
FT DOMAIN 1449 1462
FT DOMAIN 1445 1562
FT DOMAIN 1445 1480
FT REPEAT 1481 1522
FT REPEAT 1523 1562
FT REPEAT 1565 2075
FT REPEAT 1865 1910
FT REPEAT 1912 1942
FT REPEAT 1944 1975
FT REPEAT 1978 2009
FT REPEAT 2011 2042
FT REPEAT 2044 2075
FT DISULFID 24 37
FT DISULFID 31 46
FT DISULFID 63 74

```

```

POTENTIAL.
NEUROGENIC LOCUS NOTCH HOMOLOG PROTEIN 1.
EXTRACELLULAR (POTENTIAL).
POTENTIAL.
CYTOPLASMIC (POTENTIAL).
36 X EGF-TYPE REPEATS.
EGF-LIKE 1.
EGF-LIKE 2.
EGF-LIKE 3.
EGF-LIKE 4.
EGF-LIKE 5.
EGF-LIKE 6.
EGF-LIKE 7.
EGF-LIKE 8.
EGF-LIKE 9.
EGF-LIKE 10.
EGF-LIKE 11.
EGF-LIKE 12.
EGF-LIKE 13.
EGF-LIKE 14.
EGF-LIKE 15.
EGF-LIKE 16.
EGF-LIKE 17.
EGF-LIKE 18.
EGF-LIKE 19.
EGF-LIKE 20.
EGF-LIKE 21.
EGF-LIKE 22.
EGF-LIKE 23.
EGF-LIKE 24.
EGF-LIKE 25.
EGF-LIKE 26.
EGF-LIKE 27.
EGF-LIKE 28.
EGF-LIKE 29.
EGF-LIKE 30.
EGF-LIKE 31.
EGF-LIKE 32.
EGF-LIKE 33.
EGF-LIKE 34.
EGF-LIKE 35.
EGF-LIKE 36.
CYS-RICH.
3 X LIN/NOTCH REPEATS.
LIN/NOTCH 1.
LIN/NOTCH 2.
LIN/NOTCH 3.
6 X ANK MOTIF REPEATS.
ANK MOTIF 1.
ANK MOTIF 2.
ANK MOTIF 3.
ANK MOTIF 4.
ANK MOTIF 5.
ANK MOTIF 6.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.

```

```

FT DISULFID 68 87 BY SIMILARITY.
FT DISULFID 89 98 BY SIMILARITY.
FT DISULFID 106 117 BY SIMILARITY.
FT DISULFID 111 127 BY SIMILARITY.
FT DISULFID 129 138 BY SIMILARITY.
FT DISULFID 144 155 BY SIMILARITY.
FT DISULFID 149 164 BY SIMILARITY.
FT DISULFID 166 175 BY SIMILARITY.
FT DISULFID 182 195 BY SIMILARITY.
FT DISULFID 189 204 BY SIMILARITY.
FT DISULFID 206 215 BY SIMILARITY.
FT DISULFID 222 233 BY SIMILARITY.
FT DISULFID 227 243 BY SIMILARITY.
FT DISULFID 245 254 BY SIMILARITY.
FT DISULFID 261 272 BY SIMILARITY.
FT DISULFID 266 281 BY SIMILARITY.
FT DISULFID 283 292 BY SIMILARITY.
FT DISULFID 299 312 BY SIMILARITY.
FT DISULFID 306 321 BY SIMILARITY.
FT DISULFID 323 332 BY SIMILARITY.
FT DISULFID 339 350 BY SIMILARITY.
FT DISULFID 344 359 BY SIMILARITY.
FT DISULFID 361 370 BY SIMILARITY.
FT DISULFID 376 387 BY SIMILARITY.
FT DISULFID 381 398 BY SIMILARITY.
FT DISULFID 400 409 BY SIMILARITY.
FT DISULFID 416 429 BY SIMILARITY.
FT DISULFID 423 438 BY SIMILARITY.
FT DISULFID 440 449 BY SIMILARITY.
FT DISULFID 456 467 BY SIMILARITY.
FT DISULFID 461 476 BY SIMILARITY.
FT DISULFID 478 487 BY SIMILARITY.
FT DISULFID 494 505 BY SIMILARITY.
FT DISULFID 499 514 BY SIMILARITY.
FT DISULFID 516 525 BY SIMILARITY.
FT DISULFID 532 543 BY SIMILARITY.
FT DISULFID 537 552 BY SIMILARITY.
FT DISULFID 554 563 BY SIMILARITY.
FT DISULFID 570 580 BY SIMILARITY.
FT DISULFID 575 589 BY SIMILARITY.
FT DISULFID 591 600 BY SIMILARITY.
FT DISULFID 607 618 BY SIMILARITY.
FT DISULFID 612 627 BY SIMILARITY.
FT DISULFID 629 638 BY SIMILARITY.
FT DISULFID 645 654 BY SIMILARITY.
FT DISULFID 650 665 BY SIMILARITY.
FT DISULFID 666 675 BY SIMILARITY.
FT DISULFID 682 693 BY SIMILARITY.
FT DISULFID 687 702 BY SIMILARITY.
FT DISULFID 704 713 BY SIMILARITY.
FT DISULFID 720 730 BY SIMILARITY.
FT DISULFID 725 739 BY SIMILARITY.
FT DISULFID 741 750 BY SIMILARITY.
FT DISULFID 757 768 BY SIMILARITY.
FT DISULFID 762 777 BY SIMILARITY.
FT DISULFID 779 788 BY SIMILARITY.
FT DISULFID 795 806 BY SIMILARITY.
FT DISULFID 800 815 BY SIMILARITY.
FT DISULFID 817 826 BY SIMILARITY.
FT DISULFID 833 844 BY SIMILARITY.
FT DISULFID 838 855 BY SIMILARITY.
FT DISULFID 857 866 BY SIMILARITY.
FT DISULFID 873 884 BY SIMILARITY.
FT DISULFID 878 893 BY SIMILARITY.
FT DISULFID 895 904 BY SIMILARITY.
FT DISULFID 911 922 BY SIMILARITY.
FT DISULFID 916 931 BY SIMILARITY.
FT DISULFID 933 942 BY SIMILARITY.
FT DISULFID 987 998 BY SIMILARITY.
FT DISULFID 992 1007 BY SIMILARITY.
FT DISULFID 1009 1018 BY SIMILARITY.
FT DISULFID 1025 1036 BY SIMILARITY.
FT DISULFID 1030 1045 BY SIMILARITY.

```

```

FT DISULFID 1047 1056 BY SIMILARITY.
FT DISULFID 1063 1074 BY SIMILARITY.
FT DISULFID 1068 1083 BY SIMILARITY.
FT DISULFID 1085 1094 BY SIMILARITY.
FT DISULFID 1101 1122 BY SIMILARITY.

Query Match 7.7%: Score 150; DB 1; Length 2531;
Best Local Similarity 21.6%: Pred. No. 0.0002;
Matches 90; Conservative 31; Mismatches 135; Indels 160; Gaps 26;

QY 21 CECPPS--PPRCPLGVS-----LITDGCBC-----CKMCAQQLDNGCTEAAI-CD 62
DB 164 CRCPPGFHPTRODYNECSQNPGLCRHGCHNETGSTRACACATHTGPHLELPVPCS 223
QY 63 P---HNGLYCDYSGDR-----PRYALGVCAQVY-----GVGCVLDGVR-YN-- 99
DB 224 PSPCGATGCRPTGTDTHTHCACLPFGAGNGCEENVDGCGNNCKNGAGCV-DGVMTYNCR 282
QY 100 -----NGQ-----SQPN-CK-----YKTCIDGAVG-----CTP 123
DB 283 CPPEVTGQCTEDVDDECQIMPNAQNAQTCHNTHGVCNVCYNGMTGEDCSRNIDCASA 342
QY 124 LCL-----RVRRPRLMCPHPR-----RVSIPIHCCEDQWCEDDAKRPRTAPRT 168
DB 343 ACPGATGCHDRVASTFCECPHGRTGLCHLKHACISNPNCSNCDTPNVNGKRICTCPS 402
QY 169 GAFDAVGEVEAMHRCIAYTSPMSPCS--TSCGLGVS-----TRISVNAQCMP- 215
DB 403 G-----YTGP--ACSDVDDECDDLGNANCEHAGKCLNTLGSFEQCILOG 443
QY 216 -----EDESRLCNLRPCVDYHTLTKAGKCLAVYQPPASNNFTLACGISSTRSQPKYC 269
DB 444 YTGPGCEIDVNVNCISNPNCDATCLOJGE-----FQ-----CICMPGREGVYC 487
QY 270 GV-----CMDNRCCIPYKSKTIDVSFOCPDGGFSGROYLMINACNCLSNCRN 316
DB 488 EINTBCASSPCLHNGHCHM--DKITHFQCCPKFGNGLCOYDDEC-ASTPKCN 539

RESULT 15
VWF_CANPA STANDARD; PRT: 2813 AA.
ID 028295;
AC 01-NOV-1997 (rel. 35, Created)
DT 01-NOV-1997 (rel. 35, Last sequence update)
DR 01-OCT-2000 (rel. 40, Last annotation update)
DE VON WILLEBRAND FACTOR PRECURSOR (VWF).
GN F8VWF OR VWF.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
[1]
RA Stoy S.J., Shubuya H., Noneman D.J., Holzhauer J., Mohammed I.H.,
RL Johnson G.S.,
Submitted (Aug-1996) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: IMPORTANT IN THE MAINTENANCE OF HOMEOSTASIS, IT
CC PARTICIPATES IN PLATELET-VESSEL WALL INTERACTIONS BY FORMING A
CC NONCOVALENT COMPLEX WITH COAGULATION FACTOR VIII AT THE SITE OF
CC VASCULAR INJURY (BY SIMILARITY).
CC -!- SUBUNIT: MULTIMERIC (BY SIMILARITY).
CC -!- TISSUE SPECIFICITY: BLOOD (BY SIMILARITY).
CC -!- PTM: ALL CYSTEINE RESIDUES ARE INVOLVED IN INTRACHAIN OR
CC INTERCHAIN DISULFIDE BONDS (BY SIMILARITY).
CC -!- SIMILARITY: CONTAINS 3 VWF-C DOMAINS.
CC -!- SIMILARITY: CONTAINS A C-TERMINAL CYSTEINE KNOT-LIKE DOMAIN (CTCK).
CC -!- SIMILARITY: SOME, TO SILKWORM HEMOCYTIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

```

CC	entitles requires a license agreement (See <a href="http://www.jsb-sib.ch/announce/">http://www.jsb-sib.ch/announce/</a>			
CC	or send an email to <a href="mailto:license@jsb-sib.ch">license@jsb-sib.ch</a> )			
CC	-----			
DR	EMBL; L76227; AAB05549.1; -.			
DR	HSSP; P04275; IAT2.			
DR	INTERPRO; IPRO000359; -.			
DR	INTERPRO; IPRO001007; -.			
DR	INTERPRO; IPRO001846; -.			
DR	INTERPRO; IPRO02035; -.			
DR	INTERPRO; IPRO02919; -.			
DR	PFAM; PF000007; Cys_knot; 1.			
DR	PFAM; PF01826; TIL; 4.			
DR	PFAM; PF00093; vwa; 3.			
DR	PFAM; PF00093; vwc; 3.			
DR	PFAM; PF00094; vwd; 4.			
DR	PRINTS; PRO0453; VWFADOMAIN.			
DR	PROSITE; PS01185; CTCK_1; 1.			
DR	PROSITE; PS01225; CTCK_2; 1.			
DR	PROSITE; PS01208; VMFC; 3.			
KW	Blood coagulation; Platelet; Glycoprotein; Extracellular matrix;			
KW	Plasma; Endothelial cell; Repeat; Cell adhesion; signal.			
FT	SIGNAL	1	22	BY SIMILARITY.
FT	PROPEP	23	763	DL/D2 (BY SIMILARITY).
FT	CHAIN	764	2813	VON WILLEBRAND FACTOR.
FT	DOMAIN	764	787	AMINO-TERMINAL.
FT	DOMAIN	788	833	E1.
FT	DOMAIN	826	853	CX.
FT	DOMAIN	842	1130	D3.
FT	DOMAIN	1260	1479	A1.
FT	DOMAIN	1480	1672	A2.
FT	DOMAIN	1673	1874	A3.
FT	DOMAIN	1934	2203	D4.
FT	DOMAIN	2216	2261	E2.
FT	DOMAIN	2255	2326	VMFC 1.
FT	DOMAIN	2429	2495	VMFC 2.
FT	DOMAIN	2580	2650	VMFC 3.
FT	DOMAIN	2724	2812	CTCK.
FT	SITE	2507	2509	CELL ATTACHMENT SITE.
FT	DISULFID	767	808	BY SIMILARITY.
FT	DISULFID	776	804	BY SIMILARITY.
FT	DISULFID	810	821	BY SIMILARITY.
FT	DISULFID	867	996	BY SIMILARITY.
FT	DISULFID	889	1031	BY SIMILARITY.
FT	DISULFID	898	993	BY SIMILARITY.
FT	DISULFID	914	921	BY SIMILARITY.
FT	DISULFID	1060	1084	BY SIMILARITY.
FT	DISULFID	1071	1111	BY SIMILARITY.
FT	DISULFID	1089	1091	BY SIMILARITY.
FT	DISULFID	1153	1165	BY SIMILARITY.
FT	DISULFID	1149	1169	BY SIMILARITY.
FT	DISULFID	1126	1130	BY SIMILARITY.
FT	DISULFID	1196	1199	BY SIMILARITY.
FT	DISULFID	1234	1237	BY SIMILARITY.
FT	DISULFID	1272	1458	BY SIMILARITY.
FT	DISULFID	1669	1670	BY SIMILARITY.
FT	DISULFID	1686	1872	BY SIMILARITY.
FT	DISULFID	1879	1904	BY SIMILARITY.
FT	DISULFID	1899	1940	OR 1942 (BY SIMILARITY).
FT	DISULFID	1972	2123	BY SIMILARITY.
FT	DISULFID	1950	2085	BY SIMILARITY.
FT	DISULFID	1927	2088	BY SIMILARITY.
FT	DISULFID	1993	2001	BY SIMILARITY.
FT	DISULFID	2724	2774	BY SIMILARITY.
FT	DISULFID	2739	2788	BY SIMILARITY.
FT	DISULFID	2750	2804	BY SIMILARITY.
FT	DISULFID	2754	2806	BY SIMILARITY.
FT	DISULFID	?	2811	BY SIMILARITY.
FT	CARBOHYD	?	99	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	156	156	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	211	211	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	666	666	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	857	857	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	1231	1231	N-LINKED (GLCNAC. . .) (POTENTIAL).

FT	CARBOHYD	1515	1515	N-LINKED (GLCNAC. . .)	(POTENTIAL)
FT	CARBOHYD	1574	1574	N-LINKED (GLCNAC. . .)	(POTENTIAL)
FT	CARBOHYD	2223	2223	N-LINKED (GLCNAC. . .)	(POTENTIAL)
FT	CARBOHYD	2290	2290	N-LINKED (GLCNAC. . .)	(POTENTIAL)
FT	CARBOHYD	2357	2357	N-LINKED (GLCNAC. . .)	(POTENTIAL)
FT	CARBOHYD	2400	2400	N-LINKED (GLCNAC. . .)	(POTENTIAL)
FT	CARBOHYD	2546	2546	N-LINKED (GLCNAC. . .)	(POTENTIAL)
FT	CARBOHYD	2585	2585	N-LINKED (GLCNAC. . .)	(POTENTIAL)
FT	CARBOHYD	2790	2790	N-LINKED (GLCNAC. . .)	(POTENTIAL)
FT	SITE	531	533	CELL ATTACHMENT SITE (POTENTIAL).	
FT	SITE	698	700	CELL ATTACHMENT SITE (POTENTIAL).	
FT	SITE	2507	2509	CELL ATTACHMENT SITE (POTENTIAL).	
50	SEQUENCE	2813	309730	AA: 03D28D6457718B95	CRC64:

Query Match	7.7%;	Score 150;	DB 1;	Length 2813;
Best Local Similarity	23.1%;	Pred. No. 0.00021;		
Matches	92;	Conservative 48;	Mismatches 118;	Indels 140;
				Gaps 29

OY	13	RPOCKMPCEBPSPRPHCPGLVSLIHDGEE--CCKKK---AOQLDNCTEALICDPHNGL	67
Dd	2195	RANFC--AMSCRP-----SLYNHCBEHCPRCLBESNTSSCGDDPSEGCFCPPNOVM	2243
OY	68	Y--CDYSGBDRPRYAIGCAQVVGVGCVLDGVARYNNGOSFOD--NCRNYCTCIDG-AVG	120
Dd	2244	LEGGC-----VPREA---CTQCISE---DGVRHQLEFWAHQPCOI-CTCLSGRKVN	2290
OY	121	CT-----PLC----LRVR-----PRLMCPHRPRVST-	143
Dd	2291	CTLORPETARAPITGPCPEVALRIROMAECCSREVECVSDLVSCDLPRPVRCEDGIOMTLTN	2350
OY	144	PBGHCEDQWVC-----EDDKR-----PKRT-APRODGAERADNGEVAHNRRNICATVTSWS	192
Dd	2351	PGEBCRPMTACAKKDEBCARRSPSCRPHHTLARKTQCCD---EYEC-ACMCNVST----	2402
OY	193	PCTSTCGIG-VSTRISN-----VNAOCMPDEOESHLCMLR-----PCDYDIHTLI	235
Dd	2403	---VSCPGLGYIASAVTNDGCCCTTTTCFRPD--KVCHNRITITPYOGFMENBADVOCTIDL	2456
OY	236	KAGKKCLAVTYOPEAISMNFLLAGCISTRSY---ORKYGVACIMDNKCCIPRXS-----	283
Dd	2457	EDSVMGJLRVAQ--CSOKRPCEDNCLSGFTVYLHGEGCCGGRCLPSACEVUIGVSBRDAQSHW	2514
OY	284	KTIIVSFQCCPRDGLGSROYLMINACCNLSCRPNPIEF	321
Dd	2515	KNVGSHMASPD-----NPCLINECVARKKEVF	2541

Search completed: February 16, 2001, 16:15:29  
Job time: 137 sec





GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

## OM protein - protein search, using sw model

Run on: February 16, 2001, 16:12:49 ; Search time 27.01 Seconds

(without alignments)  
1453.708 Million cell updates/sec

Title: US-09-325-019-2

Perfect score: 1937

Sequence: 1 DFTAPLEDDTSRRPQCKMP.....NPNDIFADESYPPDESEIAN 335

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 374700 seqs, 117207915 residues

Total number of hits satisfying chosen parameters: 374700

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 08

Maximum Match 1008

Listing first 45 summaries

Database : SPTREMBL.15.\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phase:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp Vertebrate:\*  
14: sp\_unclassified:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1937	100.0	367	4	095388 homo sapien
2	1675	86.5	367	11	054775 mus musculu
3	852.5	44.0	347	13	09PT80 O9p180 notophthal
4	844	43.6	354	4	095389 homo sapien
5	836.5	43.2	347	11	09RIE9 O9rie9 rattus norv
6	835.5	43.1	347	11	09WVSI O9wvs1 rattus norv
7	830	42.8	343	13	042607 O42607 xenopus lae
8	827.5	42.7	349	6	097765 O97765 sus scrofa
9	815.5	42.1	331	4	095958 O95958 homo sapien
10	779.5	40.2	379	11	09WTM9 O9wtm9 rattus norv
11	762.5	39.4	381	4	043775 O43775 homo sapien
12	719.5	37.1	351	11	09QZ05 O9qz05 rattus norv
13	555.5	28.7	334	4	09UID7 O9uid7 homo sapien
14	498	25.7	251	11	09Z0G4 O9z0g4 mus musculu
15	490.5	25.3	250	4	076076 O76076 homo sapien
16	477.5	24.7	250	11	09JHC6 O9jhc6 rattus norv
17	462.5	23.9	176	13	09PS6 O9ps6 gallus gall
18	328.5	17.0	128	11	09R2C0 O9r2c0 rattus norv
19	305	15.7	100	4	09UDL6 O9udl6 homo sapien

20	275	14.2	113	11	09Z164 O9z164 rattus norv
21	218	11.3	74	11	09WUW4 O9wuw4 rattus norv
22	182.5	9.4	77	4	09UDE4 O9ude4 homo sapien
23	174.5	9.0	179	5	09VVK0 O9vvk0 drosophila
24	174.5	9.0	2327	13	091BG7 O91bg7 xenopus lae
25	169.5	8.8	4123	4	075851 O75851 homo sapien
26	161.5	8.3	1637	6	09XSV8 O9xsv8 bos taurus
27	161	8.3	1036	4	09NZV1 O9nzv1 homo sapien
28	160.5	8.3	430	5	026424 O26424 crassostrea
29	159	8.2	2482	6	028833 O28833 sus scrofa
30	153.5	7.9	3680	5	09VRO8 O9vro8 drosophila
31	150	7.7	2813	6	09T514 O9t514 canis faml
32	149	7.7	543	5	09VJ05 O9vj05 drosophila
33	149	7.7	620	5	09NKD8 O9nk8 drosophila
34	148.5	7.7	685	6	09TTS5 O9tts5-bos taurus
35	148	7.6	2813	6	028311 O28311 canis faml
36	146	7.5	765	5	09NLS0 O9nls0 sarcophaga
37	146	7.5	988	6	097867 O97867 sus scrofa
38	145.5	7.5	1042	4	013792 O13792 homo sapien
39	145.5	7.5	1081	4	076065 O76065 homo sapien
40	145	7.5	1028	11	09JLLO O9jll0 mus musculu
41	144	7.4	1444	5	017591 O17591 caenorhabdi
42	142.5	7.4	2843	4	09Y6R7 O9y6r7 homo sapien
43	142	7.3	1111	5	09XWD6 O9xwd6 caenorhabdi
44	141.5	7.3	1664	5	09YWO2 O9ywo2 homo sapien
45	141	7.3	845	4	090E28 O9ue28 homo sapien

## ALIGNMENTS

RESULT 1	PRELIMINARY;	PRT;	367 AA.
O95388	AC	O95388:	ID
01-MAY-1999 (TREMBLrel. 10, Created)	DT	01-MAY-1999 (TREMBLrel. 10, Last sequence update)	DT
01-JUN-2000 (TREMBLrel. 14, Last annotation update)	DT	01-JUN-2000 (TREMBLrel. 14, Last annotation update)	DT
CONNECTIVE TISSUE GROWTH FACTOR RELATED PROTEIN WISP-1.	DE	CONNECTIVE TISSUE GROWTH FACTOR RELATED PROTEIN WISP-1.	DE
WISPI.	GN	WISPI.	GN
Homo sapiens (Human).	OS	Homo sapiens (Human).	OS
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;	OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;	OC
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	OC
NCBI_TaxID=9606;	OX	NCBI_TaxID=9606;	OX
[1]	RN	[1]	RN
SEQUENCE FROM N.A.	RP	SEQUENCE FROM N.A.	RP
TISSUE=ADULT LUNG, AND FETAL KIDNEY;	RC	TISSUE=ADULT LUNG, AND FETAL KIDNEY;	RC
MEDLINE=99061933; PubMed=9843955;	RX	MEDLINE=99061933; PubMed=9843955;	RX
Pennica D., Swanson T.A., Welsh J.W., Roy M.A., Lawrence D.A., Lee J.,	RA	Pennica D., Swanson T.A., Welsh J.W., Roy M.A., Lawrence D.A., Lee J.,	RA
Brush J., Taneyhill L.A., Deuel B., Lew M., Watanabe C., Cohen R.L.,	RA	Brush J., Taneyhill L.A., Deuel B., Lew M., Watanabe C., Cohen R.L.,	RA
Melham M.F., Finley G.G., Quirke P., Goddard A.D., Hillan K.J.,	RA	Melham M.F., Finley G.G., Quirke P., Goddard A.D., Hillan K.J.,	RA
Gurney A.L., Botstein D., Levine A.J.;	RA	Gurney A.L., Botstein D., Levine A.J.;	RA
"WISP genes are members of the connective tissue growth factor family	RT	"WISP genes are members of the connective tissue growth factor family	RT
that are up-regulated in wnt-1-transformed cells and aberrantly	RT	that are up-regulated in wnt-1-transformed cells and aberrantly	RT
expressed in human colon tumors."	RT	expressed in human colon tumors."	RT
Proc. Natl. Acad. Sci. U.S.A. 95:14717-14722(1998).	RT	Proc. Natl. Acad. Sci. U.S.A. 95:14717-14722(1998).	RT
[2]	RT	[2]	RT
SEQUENCE FROM N.A.	RP	SEQUENCE FROM N.A.	RP
Bleichschmidt K., Kalaydjieva L., Goodman R., Gresham D., Baas F.,	RA	Bleichschmidt K., Kalaydjieva L., Goodman R., Gresham D., Baas F.,	RA
Jonge R.d., Schilhabel M., Schattevov R., Dette M., Menzel U.,	RA	Jonge R.d., Schilhabel M., Schattevov R., Dette M., Menzel U.,	RA
Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.	RA	Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.	RA
EMBL; AF100779; AAC96321.1;	DR	EMBL; AF100779; AAC96321.1;	DR
EMBL; AF192304; AAF22341.1;	DR	EMBL; AF192304; AAF22341.1;	DR
INTERPRO: IPR000359;	DR	INTERPRO: IPR000359;	DR
INTERPRO: IPR000867;	DR	INTERPRO: IPR000867;	DR
INTERPRO: IPR000884;	DR	INTERPRO: IPR000884;	DR
INTERPRO: IPR001007;	DR	INTERPRO: IPR001007;	DR
PFAM: PF00007; Cys_knot; 1.	DR	PFAM: PF00007; Cys_knot; 1.	DR
PFAM: PF00090; tsp_1; 1.	DR	PFAM: PF00090; tsp_1; 1.	DR
PFAM: PF00093; vwc_1.	DR	PFAM: PF00093; vwc_1.	DR
PFAM: PF00219; IGFBP; 1.	DR	PFAM: PF00219; IGFBP; 1.	DR
PROSITE; PS01185; CTCK_1; 1.	DR	PROSITE; PS01185; CTCK_1; 1.	DR

DR PROSITE: PS01208; VMEC; UNKNOWN\_1.  
DR PROSITE: PS01225; CTCR\_2; 1.  
SQ SEQUENCE 367 AA; 40331 MW; 9F29CA94D69C0502 CRC64;

Query Match 100.0%; Score 1937; DB 4; Length 367;  
Best Local Similarity 100.0%; Pred. No. 2,1e-183;  
Matches 335; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DTPPALEDTSSRPOCKWPCCEPPSPRCPLGVSLLITDCECKCKMAQOOLGDNCTEAAT 60  
DB 33 DTPPALEDTSSRPOCKWPCCEPPSPRCPLGVSLLITDCECKCKMAQOOLGDNCTEAAT 92  
QY 61 CDPHGLCYDYSDDRRPRYAIGVCAQVYGVCLDGYRNNNGSFQPCNKYNTCTIDGANG 120  
DB 93 CDPHGLCYDYSDDRRPRYAIGVCAQVYGVCLDGYRNNNGSFQPCNKYNTCTIDGANG 152  
QY 121 CTPCLRVPRRLMCPHPRVRSIPGHCCEQWVCEDDAKPRRTAPRDTGAFDAVGEYEA 180  
DB 153 CTPCLRVPRRLMCPHPRVRSIPGHCCEQWVCEDDAKPRRTAPRDTGAFDAVGEYEA 212  
QY 181 HNCIAVTSWSPSCSTSCGLGVSTRISNVNAOCMPQESRLCNLRPCDDVDIHLINAKKK 240  
DB 213 HNCIAVTSWSPSCSTSCGLGVSTRISNVNAOCMPQESRLCNLRPCDDVDIHLINAKKK 272  
QY 241 CLAVOPEASNMFTLAGCISTRSYOPKYGVCMDNRCCIPYKSKTIDVFPQCPDGLGFSR 300  
DB 273 CLAVOPEASNMFTLAGCISTRSYOPKYGVCMDNRCCIPYKSKTIDVFPQCPDGLGFSR 332  
QY 301 QVLMINACFCNLSCRNPNDIFADLESYPDFSEIAN 335  
DB 333 QVLMINACFCNLSCRNPNDIFADLESYPDFSEIAN 367

RESULT 2  
ID 054775 PRELIMINARY: PRT: 367 AA.

AC 054775;  
DT 01-JUN-1998 (TREMBLrel. 06, Created)  
DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)  
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)  
DE ELM1.  
GN ELM1 OR WISPL.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=HEN.  
RX MEDLINE=98119879; PubMed=9449709;  
RA Haselilmo Y., Shindo-okada N., Tanl M., Nagamachi Y., Takeuchi K.,  
RA Shiroishi T., Toma H., Yokota J.;  
RT "Expression of the ELM1 gene, a novel gene of the CCN (connective  
RT tissue growth factor, Cyr61/Ce10, and neuroblastoma overexpressed  
RT gene) family, suppresses in vivo tumor growth and metastasis of K-1735  
RT murine melanoma cells.";  
RL J. Exp. Med. 187:289-296(1998).  
RN [2]  
RN SEQUENCE FROM N.A.  
RC TISSUE=MAMMARY;  
RX MEDLINE=99061933; PubMed=9843955;  
RA Pennica D., Swanson T.A., Welsh J.W., Roy M.A., Lawrence D.A., Lee J.,  
RA Brush J., Taneyhill L.A., Deuel B., Lew M., Watanabe C., Cohen R.L.,  
RA Melham M.F., Finley G.G., Quirke P., Goddard A.D., Hillan K.J.,  
RA Guney A.L., Botstein D., Levine A.J.;  
RT "WSP genes are members of the connective tissue growth factor family  
RT that are up-regulated in wt1-transfected cells and aberrantly  
RT expressed in human colon tumors.";  
RL Proc. Natl. Acad. Sci. U.S.A. 95:14717-14722(1998).  
DR EMBL; AB004873; BAA24949.1; -  
DR EMBL; AF100777; AAC96319.1; -  
DR MGD; MGI:1197008; Wispl.

DR INTERPRO: IPR000359; -  
DR INTERPRO: IPR000867; -  
DR INTERPRO: IPR000884; -  
DR INTERPRO: IPR001007; -  
DR PEAM; PF00090; tap.1; 1.  
DR PEAM; PF00093; vwc.1; 1.  
DR PEAM; PF00219; IGFBP.1.  
DR PROSITE: PS01185; CTCR\_1; 1.  
DR PROSITE: PS01208; VMEC; UNKNOWN\_1.  
DR PROSITE: PS01225; CTCR\_2; 1.  
SQ SEQUENCE 367 AA; 40702 MW; 3B7C0569EFA5E96 CRC64;

Query Match 86.5%; Score 1675; DB 11; Length 367;  
Best Local Similarity 85.3%; Pred. No. 1.6e-157;  
Matches 285; Conservative 20; Mismatches 29; Indels 0; Gaps 0;

QY 2 FHPAPLEDTSSRPOCKWPCCEPPSPRCPLGVSLLITDCECKCKMAQOOLGDNCTEAAT 61  
DB 34 FHPAPLEDTSSRPOCKWPCCEPPSPRCPLGVSLLITDCECKCKMAQOOLGDNCTEAAT 93  
QY 62 DPHRGLCYDYSDDRRPRYAIGVCAQVYGVCLDGYRNNNGSFQPCNKYNTCTIDGANG 121  
DB 94 DPHRGLCYDYSDDRRPRYAIGVCAQVYGVCLDGYRNNNGSFQPCNKYNTCTIDGANG 153  
QY 122 TPLCLRVPRRLMCPHPRVRSIPGHCCEQWVCEDDAKPRRTAPRDTGAFDAVGEYEA 181  
DB 154 TPLCLRVPRRLMCPHPRVRSIPGHCCEQWVCEDDAKPRRTAPRDTGAFDAVGEYEA 213  
QY 182 RNCIAVTSWSPSCSTSCGLGVSTRISNVNAOCMPQESRLCNLRPCDDVDIHLINAKKK 241  
DB 214 RNCIAVTSWSPSCSTSCGLGVSTRISNVNAOCMPQESRLCNLRPCDDVDIHLINAKKK 273  
QY 242 LAVOPEASNMFTLAGCISTRSYOPKYGVCMDNRCCIPYKSKTIDVFPQCPDGLGFSR 301  
DB 274 LAVOPEASNMFTLAGCISTRSYOPKYGVCMDNRCCIPYKSKTIDVFPQCPDGLGFSR 333  
QY 302 QVLMINACFCNLSCRNPNDIFADLESYPDFSEIAN 335  
DB 334 QVLMINACFCNLSCRNPNDIFADLESYPDFSEIAN 367

RESULT 3  
ID 09PR80 PRELIMINARY: PRT: 347 AA.

AC 09PR80;  
DT 01-MAY-2000 (TREMBLrel. 13, Created)  
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)  
DE CONNECTIVE TISSUE GROWTH FACTOR.  
GN CTGF.  
OS Notoththalmus viridescens (Eastern newt) (Triturus viridescens).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Amphibia; Batrachia; Caudata; Salamandroidae; Salamandridae;  
OC Notoththalmus.  
OX NCBI\_TaxID=8316;  
RN [1]  
RN SEQUENCE FROM N.A.  
RC TISSUE=FORELIMB BLASTEMA;  
RA Gates P.B.;  
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RN SEQUENCE FROM N.A.  
RC TISSUE=FORELIMB BLASTEMA;  
RA Cash D.E., Gates P.B., Imokawa Y., Brookes J.P.;  
RT "Identification of new connective tissue growth factor as a target of  
RT retinoid regulation in limb blastema cells.";  
RL Gene 222:119-124(1998).  
DR EMBL; AJ271167; CAB65965.1; -  
DR INTERPRO: IPR000359; -  
DR INTERPRO: IPR000867; -  
DR INTERPRO: IPR000884; -  
DR INTERPRO: IPR001007; -

```

DR PFAM: PF00007: Cys_knot; 1.
DR PFAM: PF00090: tsp_1; 1.
DR PFAM: PF00093: wgc; 1.
DR PFAM: PF00219; IGFBP; 1.
DR PROSITE: PS00222; IGF_BINDING; 1.
DR PROSITE: PS01185; CTCK_1; 1.
DR PROSITE: PS01208; WFGC; UNKNOWN; 1.
SQ SEQUENCE 347 AA: 38098 MW: 387E239F27672C1 CRC64:

Query Match 44.0%; Score 852.5; DB 13; Length 347;
Best Local Similarity 44.5%; Pred. No.2.5e-76;
Matches 149; Conservative 53; Mismatches 102; Indels 31; Gaps 5

OY 15 QPCMKPCPCPSPPRCPLGVLITIDGCECCCKMAQOGLGDNCTEAALICDPHRLGYCDYSGD 74
Db 26 QDCSECCRCPPKPPCCPAGTSLVMDGCCCKCAKOLGELCTEKVDCPDHRLGFCDF -GS 84
OY 75 RPRVIGCAQVGVGCVLDGVRYNNNGSFQDNCKYKNCIDGAVGCPFLC-LRRRPPPL 133
Db 85 RNNKRGICTAKADGAPCVFVGWVYSGESFQSSCKYQCCIDGCGVPLCGVDVRLPSP 144
OY 134 WCPHRRVRSIPGHCEOWVCEDDARPPKRTAPRDGAFDAGEVEAMH----- 182
Db 145 DCPFRRRKQKLGCCCEEWCDQPKFQTR-----VGPALAVYKQETTYGPDSSL 192
OY 183 ---NCIAVTSPPWPCSTSCGLGVSTRISNVNAQCWPEQESRLCNLRPCDVIDHTLIKAK 239
Db 193 MRANLVGTTEMSASCKTGMGISIRVINDNMCKLERKQSRICAMVPCGADLENNIKKK 252
OY 240 KCLAVYQEPASNNFLAGCISTRSYQPKYGVGCMNRCCTPYKSKTIDVSFOCPDGLGS 299
Db 253 KCIIRPKISKIPKFKELSGCTSVKYRAKFCGCTGRCCTPHRTATLPLPEFKCPDGEYVK 312
OY 300 ROVLIMNACFCNLSCRNPNDIFADL---ESYPDFS 331
Db 313 KMMFIKTCACHNCPGDNDIFESMYRRKMGDMA 347

RESULT 4
O95389 PRELIMINARY: PRT: 354 AA.
AC O95389;
DT 01-MAY-1999 (TREMBlrel. 10, Created)
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
DT 01-MAY-2000 (TREMBlrel. 13, Last annotation update)
DE CONNECTIVE TISSUE GROWTH FACTOR RELATED PROTEIN WISP-3.
GN WISP3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=ADULT BONE MARROW, FETAL KIDNEY;
RX MEDLINE=99061933; PubMed=9843955;
RA Penhita D., Swanson T.A., Welsh J.W., Roy M.A., Lawrence D.A., Lee J.,
RA Brush J., Taneyhill L.A., Denel B., Lew M., Watanabe C., Cohen R.L.,
RA Melham M.F., Finley G.G., Quikre P., Goddard A.D., Hillan K.J.,
RA Guney A.L., Botstein D., Levine A.J.;
RT "WISP genes are members of the connective tissue growth factor family
RT that are up-regulated in wnt-1-transformed cells and aberrantly
RT expressed in human colon tumors";
RL Proc. Natl. Acad. Sci. U.S.A. 95:14717-14722(1998).
DR EMBL: AF100781; AAC96323.1; -.
DR INTERPRO: IPR000359; -.
DR INTERPRO: IPR000867; -.
DR INTERPRO: IPR000884; -.
DR PFAM: PF00007: Cys_knot; 1.
DR PFAM: PF00090: tsp_1; 1.
DR PFAM: PF00219; IGFBP; 1.
DR PROSITE: PS00222; IGF_BINDING; 1.
DR PROSITE: PS01225; CTCK_2; 1.

```

[illegible]

```

QY      68  YCDYSGSRPRALIVCAQAAVGVGCVLLDGVARNNOCSOPMKCYCTIDAVGCTPLC-L 156
Dh      79  FCDP- GSPANKITGVCAKQGAAPCVFSGSVYRSESTQSSCKTQCTLDGAVGCVPLCSM 137
QY      127 RYRPRRLICPHPRRVSITDGHCCEDQVCEBDARKPRKTAIPDGTGAFDAVGEYEAHNR- 1022
Dh      138 DVRLPSDPCPPRRVRKTLPGKCEEMWQDE-----PKDR--TYVGGALAAAYRLDET 185
QY      183 -----NCIAYTSPWSVCSCTSGGLGAVSTRISVNNACQWPEQESRLNLNRPQVDIH 2322
Dh      186 FGPDPTMMRANCIQVOTIEMSAKSCSTCGMGISTRTNTNTRCRLKESGRSLDMVRPCEADLE 2455
QY      233 TLIAGAKCCLAVOPEASMNFTLAGIISTRSYOPRYGVCQMDNRCCIPYKSKTIDVSFOC 2922
Dh      246 ENIKKKKCIPTPRIAPVYKFELSGCSTVATYRAKFCGVCYCDGRCCTPHRTTLLPVEFKC 3055
QY      293 PDGLGFSROYVIMINACFCNLSCRPNIDIFDL---ESTPDS 331
Dh      306 PDGEIMKKNNMFITTCACHYCNCPEDNIDFSLVLYRKAYGMA 347

```

RESULT	6			
09WVS1				
ID	09WVS1	PRELIMINARY;	PRT;	347 AA.
DT	01-NOV-1999	(TREMBLrel. 12, Created)		
DT	01-NOV-1999	(TREMBLrel. 12, last sequence update)		
DT	01-JUN-2000	(TREMBLrel. 14, last annotation update)		
DE	CONNECTIVE TISSUE GROWTH FACTOR.			
OS	Rattus norvegicus (Rat).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.			
NCBI_TaxID=10116;				
[1]				
SEQUENCE FROM N.A.				
RA	Tezuka K., Tamatani T.;			
RT	"Rattus norvegicus connective tissue growth factor";			
RL	Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; AB023068; BA082125.1; -.			
DR	INTERPRO: IPR000359; -.			
DR	INTERPRO: IPR000867; -.			
DR	INTERPRO: IPR000884; -.			
DR	INTERPRO: IPR001007; -.			
DR	Pfam: PF00007; Cys_knot; 1.			
DR	Pfam: PF00090; tsp_1; 1.			
DR	Pfam: PF00093; vwc; 1.			
DR	Pfam: PF00219; IGFBD; 1.			
DR	PROSITE: PS00222; IGF_BINDING; 1.			
DR	PROSITE: PS01185; CTCK_1; 1.			
DR	PROSITE: PS01208; VMPC; 1.			
DR	PROSITE: PS01225; CTCK_2; 1.			
DR	SEQUENCE 347 AA; 37837 MW; 6A69511DE72FBFC CRG64;			

Query Match	44.1%	Score 835.5	DB 11	Length 347
Best Local Similarity	44.4%	Pred. No. 1,2e74		
Matches	152	Conservative	55	Mismatches 100, Indels 35, Gaps 8

  

OY	12	SRP----	QCKMPCDEP-FSPNCPGLVSLITDGCBECKMCAOGLDNCITBEAICDPHRLG	67
Db	19	TRPATGDCSAOCCQCARERAPRC	PAVSLVTDGCGCCRCACAOGLGTLCTERDPDCKHKL	78
OY	68	YCDYSGDRPRAYAGCAOVVGV	CLDYGVRNNGQSPFQNCCKRYNCTCIDGAVCTPLC-L	126
Db	79	FCDF-GSPANRRIGCAPKMDGAP	CFVGGSVYSGSGFSQSCCKYCCCTCLGAVGVCTPLCSM	137
OY	127	RVRPRRLCPHRRRYSIFGHCC	EOVMCEBDKARPKRTAPRDTGARDAYGEVEMNR----	182
Db	138	DVRLESPCCPFPRRRKTLGKCC	EBWYDC-----PKDR-----TVVGPALAAVYLEDT	185
OY	183	-----NCIATSPWSPCSTGCG	LAGVSTRISNVNAQCPEDESRCLNLRPCDDYDIH	232

```

Db      186  FGPPTMMRANCLVOTUEWMSACSTCCMGJSTPVTNNTFCRLKEQSRGLCMVRPCEDLE 2455
Oy      233  TLIRAGKRCJLAVYVPBASMNNTLGIISTSTSYOKRYGVCVMDNNCCIPYKSTIDVSFOC 2922
Db      246  ENIKKGKCCITPPIARPYVELSGCISVKTIRAKFCGVCYCTDGCCTCPHPTTLTPYEFKC 3055
Oy      293  PDGLGFSQVYIMINACTCUNISCRNPNDIFADL---ESTPDPS 331
Db      306  PDGEIMRKNMFIPTCACHYCCPDNDIDFESLYYRKRYGMA 347

```

RESULT	7	
ID	042607	PRELIMINARY; PRT: 343 AA.
AC	042607;	
DT	01-JAN-1998 (TEMBLrel. 05, Created)	
DT	01-JAN-1998 (TEMBLrel. 05, Last sequence update)	
DT	01-JUN-2000 (TEMBLrel. 14, Last annotation update)	
DE	CONNECTIVE TISSUE GROWTH FACTOR XCTGF.	
DE	Xenopus laevis (African clawed frog).	
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC	Amphibia; Batrachia; Anura; Mesobatrachia; Pipiloidea; Pipidae;	
OC	Xenopodidae; Xenopus.	
OX	NCBI_TaxID=8355;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RA	ying Z., King M.L.	
RL	Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.	
DR	EMBL: U43524; AAB67639.1; -	
DR	EMBL: U43523; AAB67638.1; -	
DR	INTERPRO: IPR000359; -	
DR	INTERPRO: IPR000867; -	
DR	INTERPRO: IPR000884; -	
DR	INTERPRO: IPR001007; -	
DR	PFAM: PF00007; Cys_Knot; 1.	
DR	PFAM: PF00090; tsp_1; 1.	
DR	PFAM: PF00093; vwc_1; 1.	
DR	PFAM: PF00219; IGFBP; 1.	
DR	PROSITE: PS01185; CTCK_1; 1.	
DR	PROSITE: PS01208; WWC; 1.	
DR	PROSITE: PS01225; CTCK_2; 1.	
QC	SEQUENCE 343 AA; 57966 MW; 93F221C5DB565A61 CRC64;	

[illegible]

RESULT 8  
ID 097765 PRELIMINARY; PRT: 349 AA.  
AC 097765;  
DT 01-MAY-1999 (Tremblrel. 10, Created)  
DT 01-MAY-1999 (Tremblrel. 10, last sequence update)  
DT 01-JUN-2000 (Tremblrel. 14, last annotation update)  
DE CONNECTIVE TISSUE GROWTH FACTOR.  
GN CTGF.  
OS Sus scrofa (Pig).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
OX NCBI\_TaxId=9823;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE-UTERUS;  
RA Harding P.A., Brigsstock D.R.;  
RT "Cloning and sequencing of a porcine connective tissue growth factor  
RT (CTGF) cDNA."  
RL Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.  
DR EMBL: U70060; AAD00174.1; -.  
DR INTERPRO: IPR000359; -.  
DR INTERPRO: IPR000867; -.  
DR INTERPRO: IPR000884; -.  
DR INTERPRO: IPR001007; -.  
DR PFAM: PF00007; Cys\_knot; 1.  
DR PFAM: PF00090; tsp\_1; 1.  
DR PFAM: PF00093; vwc\_1; 1.  
DR PFAM: PF00219; IGFBP; 1.  
DR PROSITE: PS00222; IGF\_BINDING; 1.  
DR PROSITE: PS01185; CTCK\_1; 1.  
DR PROSITE: PS01208; VMFC; 1.  
DR PROSITE: PS01225; CTCK\_2; 1.  
SQ SEQUENCE 349 AA; 37946 MW; 35AB4275AC1D4B3A CRC64;

Query Match 42.7%; Score 827.5; DB 6; Length 349;  
Best Local Similarity 44.2%; Pred. No. 7.3e-74;  
Matches 151; Conservative 53; Mismatches 103; Indels 35; Gaps 8;

QY 12 SRP---QFCMKPECEPPSPR-CPGLVSLITDGECCCKMCAQOLGDNCTEATIDDPHRL 67  
DB 21 SRPASGGDCGGCCAGGRACRACAGVSLLEGCCRLCAKHLGDLCTERAPCDPHKGL 80  
QY 68 YCYSGRPRRAIGCAQVGVGLDGVRYNNGSFOPNCKYCTCIDAGVCTPLC-L 126  
DB 81 FCFD-GSPANRKGIVCTAKDCAPCFEGTVYRSSEFSQSCYCTCLDGAVGCVPICSM 139  
QY 127 RVBPRLMCPHPRVSTPGHCEQWVCEDDAKRPRKTAAPRDTGAFFDAVGEVEMHR--- 182  
DB 140 DVRLPSPDCFPFRVVKLPSCCEWVCE-----PKD--HIVVGPALAAVRLDT 187  
QY 183 -----NCIAYTSPWPCSTSCGLGVSTRISNVNAOCWPEQESRLCNLRPCVDYDIH 232  
DB 188 FGPDPMTMRANCLVQTEWSACSKTCGMGISTRTYTNDAOSRLKQSLCVRCEADLE 247  
QY 233 TLKAGKCAVQPEASNMFTLAGCISTRSYOPKYGVCMDNRCCLPIYKSKITIDVSFOC 292  
DB 248 ENRKKGKCIKRTPKISKVFELSGCTSVKTYRAKFCGVCVTDGRCPPHRTTLPVEFKC 307  
QY 293 PDGLGFSROYLWINACFCNLSCRNPNDIFADL--ESYPPDS 331  
DB 308 PDGEVWKMSMFIKTCACHYNCPCDNDIFSLYRKMYGDMA 349

RESULT 9  
ID 095958 PRELIMINARY; PRT: 331 AA.  
AC 095958;  
DT 01-MAY-1999 (Tremblrel. 10, Created)  
DT 01-MAY-1999 (Tremblrel. 10, last sequence update)  
DT 01-MAY-2000 (Tremblrel. 13, last annotation update)  
DE DL142L7.3 (CONNECTIVE TISSUE GROWTH FACTOR (NOV, GIG) LIKE PROTEIN)

DE (FRAGMENT).  
GN DL142L7.3 OR L1BC.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
OX NCBI\_TaxId=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Tubbby B.;  
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE-MAMMARY GLAND;  
RA van Golen K.L., Davies S., Wu Z.F., Wang Y.F., Bucana C.D., Root H.,  
RA Chandrasekharappa S., Strawderman M., Elhier S.P., Merajver S.D.,  
RT "A novel putative IGF-binding, tumor suppressor protein, L1BC, and  
RT R10C GTPase, are determinants of the inflammatory breast cancer  
RT phenotype."  
RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL: Z99289; CAB16556.1; -.  
DR EMBL: AF143679; AAD31517.1; -.  
DR INTERPRO: IPR000359; -.  
DR INTERPRO: IPR000867; -.  
DR INTERPRO: IPR000884; -.  
DR PFAM: PF00007; Cys\_knot; 1.  
DR PFAM: PF00090; tsp\_1; 1.  
DR PFAM: PF00219; IGFBP; 1.  
DR PROSITE: PS00222; IGF\_BINDING; 1.  
DR PROSITE: PS01225; CTCK\_2; 1.  
FT NON\_TER 331  
SQ SEQUENCE 331 AA; 36909 MW; D109C2FDCALIDF549 CRC64;

Query Match 42.1%; Score 815.5; DB 4; Length 331;  
Best Local Similarity 43.2%; Pred. No. 1.1e-72;  
Matches 143; Conservative 53; Mismatches 98; Indels 37; Gaps 6;

QY 11 SSRQFCKMPCEPPSPRCPRLGLVSLITDGECCCKMCAQOLGDNCTEATIDDPHRLGLCD 70  
DB 20 AGRQFCHWPKCPQQRPRCPGVLVNDGGCCCKLCAKQGEICNEADLCDDPHKGLYCD 79  
QY 71 YSGDRPRRAIGCAQVGVGLDGVRYNNGSFOPNCKYCTCIDAGVCTPLCLRRP 130  
DB 80 YSVDPREYETGVCA-CASVGCFFNQHNGQVQPNPLFSCILVSGAIGCTPLFI--- 134  
QY 131 PRLMCPHPRVSTPGHCEQWVCEDDAKRPRKTAAPRDTGAFFDAVGEVEMHR--- 179  
DB 135 PKL-----AGSHC-----SGAKGKKSDQSNCSLEPDLQQLSTSYKTPPAYRNL 178  
QY 180 ---WHRNCIATSPWPCSTSCGLGVSTRISNVNAOCWPEQESRLCNLRPCVDYDIHLLIK 236  
DB 179 PLIKKRCCLVQATWTPDSRPGCGISNRYTNSNENEMREKRLCYIQPDSNLLKTIK 238  
QY 237 --AGKRCIAVQPEASNMFTLAGCISTRSYOPKYGVCMDNRCCLPIYKSKITIDVSFOCPD 294  
DB 239 IPKRGKQPTQLSKAEKFFVSGSSSTOSYKPTFCGICLDKRCCLPKSKMITTIQFDCPN 298  
QY 295 GLGFSROYLWINACFCNLSCRNPNDIFADLE 325  
DB 299 EGSFKMKMLWITSCVQCRNCRPEPDIFSELK 329

RESULT 10  
ID 09WTM9 PRELIMINARY; PRT: 379 AA.  
AC 09WTM9;  
DT 01-NOV-1999 (Tremblrel. 12, Created)  
DT 01-NOV-1999 (Tremblrel. 12, last sequence update)  
DT 01-JUN-2000 (Tremblrel. 14, last annotation update)  
DE CYR61 PRECOURSOR.  
GN CYR61.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;















```

QY      68  YCDYSGDPPRAIVCAQAAVGVGVLGVRARNNOCSOPMKCYKSTCIDAVAGCTPLC-L 126
Db      81  FCDP-GSPANKITGVCAKADAPCIPIFGYIYRSESPQSSCKYCTCIDDAVAGMPLCSM 139
QY      127  RVPRLRLMCPHRRVRSITPGHCCEBQWVCEDDAKRPRTKAPDGTGAFDAVGEVEMHR--- 182
Db      140  DVRLPSPDCPPRRVRKLPKGCCEBQWYDE-----PKDQ--TVVGGALAAYLEDT 187
QY      183  -----NCIAYTSPWSPCSTSGIGAVSTRISNNVNAACWPEOESRLNLPRQVDIH 232
Db      188  FGPDPTLRANCLVQTTBMSACSTGCGKIGSTRVTNNASCRLEKQSRLLMVRCEADLE 247
QY      233  TLIAGKCCIAVQPEASMNFTLGIACSTRSOPRYCGVGMCDNRCCIPYKSKTIDVSFOC 292
Db      248  ENIKKKGCIRTPIKSRPIKFIETLSCGTSMTKYRAKFGCVGCTDGRCTPHRTTLPLVEFK 307
QY      293  PDGLGSRQVYIMINACRCNLSCRPNIDFADL---ESTPPES 331
Db      308  PDGEVYMKNNMFIKTCACHYCCPDNDIFSLVLYRKAYGMDA 349

```

```

RESULT      2
US-08-386-680-2
Sequence 2, Application US/08386680
Patent No. 5585270
GENERAL INFORMATION:
APPLICANT: Grotenordt, Gary R.
APPLICANT: Bradham Jr., Douglas M.,
TITLE OF INVENTION: CONNECTIVE TISSUE GROWTH FACTOR
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Spensley Horn Jubs & Lubitz
STREET: 4225 Executive Square, Suite 1400
CITY: La Jolla
STATE: CA
COUNTRY: US
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/386,680
FILING DATE: 10-FEB-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/167,628
FILING DATE:
APPLICATION NUMBER: US/07/752,427
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Wetherell, Jr. Ph.D., John W.
REGISTRATION NUMBER: 31,678
REFERENCE/DOCKET NUMBER: PD-1294
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-455-5100
TELEFAX: 619-455-5110
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 349 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE type: protein
US-08-386-680-2

```

Query Match	44.9%	Score 869.5	DB 1	Length 349
Best Local Similarity	45.6%	Pred. No. 2.6e73		
Matches 156	Conservative 53	Mismatches 99	Indels 35	Gaps 8
12 SRP--GFCWCECPSPD-PPCPGLVSLITDGCCECKCAQALQDNCITEAICDPHRL	67			

Db	21	SRPVGQNCSCGCRCPDEBPAPRCAGVSLVLDGGCCRCVAKQGLGELCTERDDPHKGL	80
QY	68	YCDYSGDRPRRAIVCAQAVYGVCLDGVKRYNNQGSQDPMKKYNCTIDGAVGCTPLC-L	126
Db	81	FCDF-GSPANRKICVCAKAGAPCIPFGSTVYRSGESQSSCKYCTCLDGAVGCMPLCSM	139
QY	127	RVRPRRLMCPHRRVSTIPGHCCEQWVCEDDAKRRKRTAPRDTGAFAFVAVGEAMHR---	182
Db	140	DVRLRSPDCPPRRYKVLGKCCCEWVDE-----PKDQ--TVYGPAALAYRLDPT	187
QY	183	-----NCIAVTPSWSCSTSGELGYSTRISNVNAOCMPDEQSRCLNLRPCVDIH	232
Db	188	FGRPPTMIRANCLVQTTETEMSAQSTCGMGIISTRTVTDNNAQRLKSGRSLDMVRCEADLE	247
QY	233	TLIRAGKCLAVYPPESAMNFTLAGCISTSNRYOKRYGCGMCDNRCLIPVYSKTIIDVSFOC	292
Db	248	ENIRKKGKCIPTPKISPIKFKELSGCTSMKTYRAKFCGVCTDGHCTPHRTTLTPVEFKC	307
QY	293	PDGLGFSQVLMINACPCNLSCRPNIDFADL---ESYPPES	331
Db	308	PDGEVKNKMMFICTCAHNCPCDNDIFESLYRRKNTIGMA	349

```

RESULT 3
US-08-459-717-2
: Sequence 2, Application US/08459717
: Patent No. 5770209
: GENERAL INFORMATION:
: APPLICANT: Grotenordt, Gary R.
: APPLICANT: Bradham Jr., Douglas M.
: TITLE OF INVENTION: CONNECTIVE TISSUE GROWTH FACTOR
: NUMBER OF SEQUENCES: 2
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Spensley Horn Judas & Lubitz
: STREET: 4225 Executive Square, Suite 1400
: CITY: La Jolla
: STATE: CA
: COUNTRY: US
: ZIP: 92037
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/459,717
: FILING DATE: 02-JUN-1995
: CLASSIFICATION: 536
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/752,427
: FILING DATE: 30-AUG-1991
: ATTORNEY/AGENT INFORMATION:
: NAME: Wetherell, Jr., Ph.D., John W.
: REGISTRATION NUMBER: 31,678
: REFERENCE/DOCKET NUMBER: PD-1294
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 619-455-5100
: TELEFAX: 619-455-5110
: INFORMATION FOR SEQ ID NO: 2:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 349 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
:
US-08-459-717-2

```

Query Match	44.9%	Score 869.5;	DB 1;	Length 349;
Best Local Similarity	45.6%	Pred. No. 2,6e-73;		
Matches 156;	Conservative 53;	Mismatches 96;	Indels 35;	Gaps 8;
Oy	12 SRP---	OFCKMPCBCPPSP-PPCPDGLSVSLITTCGCECKMKCAQADIGDNCETEAATCDPHRL	67	

Db 21 SRVAVGNGSGPCRCDEPAPRCBAGVSLVLDGCGCCRCVCAKQDGLCTERDPCDPHKL 80  
QY 68 YCYSYSGDRPRYALGVCAGVAVGCVLDGVRYNNGQSFQPCCKNCTCIDAVGCTPLC-L 126  
Db 81 FCFD-GSPANRKGIVGCAKDGAPCIEFGYVRSSEFQSSCKYQCTCLDVAVGCMPLCSM 139  
QY 127 RVPRRLMCPHPRRVSLPGSCCEOWWCEDDAKPRKTPAPDGTAFDAVGEVAMHR---- 182  
Db 140 DVRLPSPDCFPFRVKLPKGCCEWVDE-----PDDQ---TYVGPALAAVRLDPT 187  
QY 183 -----NCIAYTSPMSPCSTSCGLGVSTRISVNAOCWPEQESRLCNLRPCVDIH 232  
Db 188 FGDPPTMIRANCLVQTTWMSAGSKTCGMGISTRTVNDNASCRLKQSLCMVRPCDADLE 247  
QY 233 TLKAGKCLAVQPEASNMFTLAGCISTRSYQKCYGVCMDNRCCIPYKSTIDVSEFC 292  
Db 248 ENIKKGKCIPTPKISKPIFELSGCTSMKTYRAKFCGCTDGRCPHRTTLLPVEFKC 307  
QY 293 PDGLGFSROYLMINACFCNLSCRNPNDIFADL---ESYPPDS 331  
Db 308 PDGEVMMKMMFTKTCACHYNCPCGNDIFESLYYRKMYGDMA 349  
RESULT 4  
US-08-712-302-2  
Sequence 2, Application US/08712302  
Patent No. 5783187  
GENERAL INFORMATION:  
APPLICANT: Grotendorst, Gary R.  
APPLICANT: Bradham Jr., Douglas M.  
TITLE OF INVENTION: CONNECTIVE TISSUE GROWTH FACTOR  
NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Spensley Horn Jubas & Lubitz  
STREET: 4225 Executive Square, Suite 1400  
CITY: La Jolla  
STATE: CA  
COUNTRY: US  
ZIP: 92037  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/712,302  
FILING DATE: 11-SEP-1996  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/386,680  
FILING DATE: 10-FEB-1995  
APPLICATION NUMBER: US/08/167,628  
FILING DATE:  
APPLICATION NUMBER: US/07/752,427  
ATTORNEY/AGENT INFORMATION:  
NAME: Wetherell, Jr. Ph.D., John W.  
REGISTRATION NUMBER: 31,678  
REFERENCE/DOCKET NUMBER: PD-1294  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 619-455-5100  
TELEFAX: 619-455-5110  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 349 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-712-302-2  
Query Match 44.9%; Score 869.5; DB 1; Length 349;

Best Local Similarity 45.6%; Pred. No. 2, 6e-73;  
Matches 156; Conservative 53; Mismatches 98; Indels 35; Gaps 8;  
QY 12 SRP---QFCRWPECPSP--PRCPGLVSLITDGCCECKMCAQDGNCTEAATCDPRLGL 67  
Db 21 SRVAVGNGSGPCRCDEPAPRCBAGVSLVLDGCGCCRCVCAKQDGLCTERDPCDPHKL 80  
QY 68 YCYSYSGDRPRYALGVCAGVAVGCVLDGVRYNNGQSFQPCCKNCTCIDAVGCTPLC-L 126  
Db 81 FCFD-GSPANRKGIVGCAKDGAPCIEFGYVRSSEFQSSCKYQCTCLDVAVGCMPLCSM 139  
QY 127 RVPRRLMCPHPRRVSLPGSCCEOWWCEDDAKPRKTPAPDGTAFDAVGEVAMHR---- 182  
Db 140 DVRLPSPDCFPFRVKLPKGCCEWVDE-----PDDQ---TYVGPALAAVRLDPT 187  
QY 183 -----NCIAYTSPMSPCSTSCGLGVSTRISVNAOCWPEQESRLCNLRPCVDIH 232  
Db 188 FGDPPTMIRANCLVQTTWMSAGSKTCGMGISTRTVNDNASCRLKQSLCMVRPCDADLE 247  
QY 233 TLKAGKCLAVQPEASNMFTLAGCISTRSYQKCYGVCMDNRCCIPYKSTIDVSEFC 292  
Db 248 ENIKKGKCIPTPKISKPIFELSGCTSMKTYRAKFCGCTDGRCPHRTTLLPVEFKC 307  
QY 293 PDGLGFSROYLMINACFCNLSCRNPNDIFADL---ESYPPDS 331  
Db 308 PDGEVMMKMMFTKTCACHYNCPCGNDIFESLYYRKMYGDMA 349  
RESULT 5  
US-08-880-031-2  
Sequence 2, Application US/08880031  
Patent No. 5916756  
GENERAL INFORMATION:  
APPLICANT: Grotendorst, Gary R.  
APPLICANT: Bradham Jr., Douglas M.  
TITLE OF INVENTION: CONNECTIVE TISSUE GROWTH FACTOR  
NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Spensley Horn Jubas & Lubitz  
STREET: 4225 Executive Square, Suite 1400  
CITY: La Jolla  
STATE: CA  
COUNTRY: US  
ZIP: 92037  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/880,031  
FILING DATE:  
APPLICATION NUMBER: US/08/167,628  
ATTORNEY/AGENT INFORMATION:  
NAME: Wetherell, Jr. Ph.D., John W.  
REGISTRATION NUMBER: 31,678  
REFERENCE/DOCKET NUMBER: PD-1294  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 619-455-5100  
TELEFAX: 619-455-5110  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 349 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-880-031-2  
Query Match 44.9%; Score 869.5; DB 2; Length 349;

Best Local Similarity 45.68; Pred. No. 2.6e-73;  
Matches 156; Conservative 53; Mismatches 98; Indels 35; Gaps 8;

```

QY 12 SRP---OEKKMPCCSPSP-PRCLGSLITLTDGCECCMKCAOOLGDNCTAFATLIDPRGL 67
Db 21 SRPVGQCSGPCBCDPEBPAPRAGVSLVLDGCGCCRVCAKOLGELCTTERDPDPHKL 80
QY 68 YCDPSGDRPRAIVCAQAVVSGCVLLDGVRRNNQSFQPMCKYNTCTSIDAVGCTPLC-L 126
Db 81 FCDP-GSPANKIGVCTIAKDAQPCIRFSGTYRSESFQSSCKTQCTLLDAVGMPLCSM 139
QY 127 RVPRRLMCPHRRVRSIPGHCCEDQWCEDAKRPRTKAPRDTGAFDAVGEVEMHR---- 182
Db 140 DVRLSPDPCPPRRVXKLPGKCCCEMWODE-----PMDQ---TYVGALAAVRLIEDT 187
QY 183 -----NCIAVTSPMWPCSTSGGLGVSTRISVNAACMPBQESRLCNLRPCVDIH 222
Db 188 FGDPPTMIRANCLVOTTEWMSACSTKCGMISTRTVTDNNASCRLEKQRLCMVRCEADLE 247
QY 223 TLIRAGKCLAVYPPESAMNTLTAGCISTSYQPKYGVGCMDBRCCLIPYASKTIDVSFOC 252
Db 248 ENIKGKKCIPTPKISPIKFIETELSGCSTSMKTYRAFCGVCTDGRCKCPHRTTLLPVEFKC 307
QY 293 PDGLGFSROYLIMINACPNLSCRPNIDFADL---SEYDPDS 331
Db 308 PDGEVMMKNMAFIYTCACHYCCPDNDIFSLVLYRKMYGMA 349

```

RESULT 6  
US-09-054-368-2

```

1  GENERAL INFORMATION:
2  APPLICANT: University of South Florida
3  APPLICANT: Grotenborg, Gary R
4  APPLICANT: Bradham, Jr., Douglas M.
5  TITLE OF INVENTION: CONNECTIVE TISSUE GROWTH FACTOR
6  FILE REFERENCE: 0741/003005
7  CURRENT APPLICATION NUMBER: US/09/054,368
8  CURRENT FILING DATE: 1998-04-02
9  EARLIER APPLICATION NUMBER: 08/386,660
10 EARLIER FILING DATE: 1995-02-10
11 EARLIER APPLICATION NUMBER: 08/459,717
12 EARLIER FILING DATE: 1995-06-02
13 EARLIER APPLICATION NUMBER: 08/167,628
14 EARLIER FILING DATE: 1993-12-14
15 NUMBER OF SEQ ID NOS: 9
16 SOFTWARE: FASTSEQ for Windows Version 4.0
17 SEQ ID NO 2
18 LENGTH: 349
19 TYPE: PRT
20 ORGANISM: Homo Sapiens
21 OS-09-054-368-2

```

Query Match	44.9%	Score 869.5;	DB 3;	Length 349;
Best Local Similarly	45.6%	Pred. No. 2.6e-73;		
Matches 156; Conservative	53;	Mismatches 98;	Indels 35;	Gaps 8;

```

QY 12 SRP---QFCMKRRCPCPSRP-PRCPGLYSLTITDGECCCKMAQOLGDNCTPAALICDPHRL 67
Db 21 SRPVGQMGSCGPCRCBPBPAPRCAGVSLYIDGGCCRCVCAKOLGEICTERDPCDPHKL 80
QY 68 YCDYSGDRPRALIGCAQVAVGCVLDGVRNNQGSQPMCKRYNCTCIDAVGCTPLC-L 126
Db 81 PCDE-GSPANKRIYCAKADQAPCIFGGTYVRSGSEFSOSKCYCTCIDAVAGCMPLCSM 139
QY 127 RVRPRRLMCPHRRVRVSTPGHCCEBDWVEDDAKRRKRTAPRDTGAFDAVGEVAMHR---- 132
Db 140 DVRLPSPDCPPRRRYKLPKGCCFEMWDE-----PKDQ---TVVGGPALAAYRLDET 187
QY 183 -----NCIAYTSPWSSCTSCGGLGYSTRISNNNAOCWPEQSERLCNLRCQVDIH 232

```

Db 188 FGPDPTMI RANVLQYTTETBMSAKSTCGMGISTRTVNDNNA SCLERLEKQGRILCVMRCEADLE 247

QY 233 TLIRKRGKCLAVQPEASMNFTLGGISTSBRYOKYGVGCMNDNCCIPYKSTIDVSEOC 292

Db 248 ENIKGKGCICIRTPKSKIRTELSECSKSMKTYRAKFGCVGLDGCCTPHRTTILPVEFK 307

QY 293 PDGLGFSQVITMINACFCNLSCRPNDFADL--ESTPDFS 331

Db 308 PDGEVYMKRNNMFIKTCAACHYCCPDGNDITFESLYIRKMGDGA 349

RESULT 7  
US-09-097-179-2  
; Sequence 2, Application US/09097179  
; Patent No. 6140016

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
84

Query Match	44.98	Score 869.5	DB 3	Length 349
Best Local Similarity	45.68	Pred. No. 2.6e-73		
Matches 156; Conservative	53	Mismatches 98	Indels 35	Gaps 8

```

0Y 12 SRF----QCKRNPCEBPSP-PRCPGLSVLTIDGCEBCKMCAQUGDNCTAATLDDPHKGL 67
Db 21 SRPAVGQNCSCGRCPDPBPAPRCAPGASVLYLDGGCCRCVCAKQJGELCTEEDPDDPHKGL 80
0Y 68 YCDYSGDRPRAATGCAQAVGVGCLDQVRYNMCOSFOPNCKYKNCOTCIDGAVGCTPLC-L 138
Db 81 FCDYF-GSPANKRIKICTKADQAPCLFGSTVYRSGESFQSSCKYQCTCLDGAQGMPLCSM 139
0Y 127 RVRPRPLMCPHRRVRSIDGHCCEDQVCEDDAKRPRTKAIPDTGAFDVAQVEYEAMHR--- 182

```

```

Db 140 DVLRLSPDCPPRRVKLPKCCCEWVDE-----PKDQ---TWVGPALAAVRLDPT 187
QY 183 -----NCIATYSPWSPCSTSCGLSTRTISNNACWPEOERLCLNLRPCVDIH 232
Db 188 FGDPPTMIRANCLVQTTWMSACSTCGMISTRTVNDNASCRLEKOSRLCMVRCEADLE 247
QY 233 TLKAGKCLAVYQPEASMTFLAGCISTRSYOPKCYGVCMDNRCCIPYKSKTIDVAFQC 292
Db 248 ENIKKGGKCIKRTIRKISPIFELSGCTSMKTYRAKFGCVCTDGRCCCPHRTTILPVEFKC 307
QY 293 PDGLGFSROYLWIMINACFCNLSCRNPNDIFADL---ESYPDFS 331
Db 308 PDGEVMMKMMFIKTCACHYNCPCDNDIFESLYRKMVGMA 349

```

```

RESULT 8
US-09-054-274-2
; Sequence 2, Application US/09054274
; Patent No. 6150101
; GENERAL INFORMATION:
; APPLICANT: University of South Florida
; APPLICANT: Grotendorst, Gary R.
; APPLICANT: Bradham, Jr., Douglass M.
; TITLE OF INVENTION: METHODS OF IDENTIFYING A COMPOSITION
; TITLE OF INVENTION: THAT ALTERS CONNECTIVE TISSUE GROWTH
; FILE REFERENCE: 07414/003004
; CURRENT APPLICATION NUMBER: US/09/054, 274
; EARLIER FILING DATE: 1998-04-02
; EARLIER APPLICATION NUMBER: 08/386, 680
; EARLIER FILING DATE: 1995-02-10
; EARLIER APPLICATION NUMBER: 08/459, 717
; EARLIER FILING DATE: 1995-06-02
; EARLIER APPLICATION NUMBER: 08/167, 628
; EARLIER FILING DATE: 1993-12-14
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 349
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-054-274-2

```

```

Query Match 44.9%; Score 869.5; DB 3; Length 349;
Best Local Similarity 45.6%; Pred. No. 2.6e-73;
Matches 156; Conservative 53; Mismatches 98; Indels 35; Gaps 8;

QY 12 SRP---QFCWPCPCPPSP-PRCPGLVSLTTDGCCECKMCAQULGDMCTEATCDPRL 67
Db 21 SRAVGNCGSGPCPCPEPAPRCAGVSLVDGCGCCRVCAKOLGELCTERDPCDPHKL 80
QY 68 YCDYSGDRPYAIGVCAQVGVGCVLDGVYRNNQSFOPKCYNCCTCIDGAVGCTPLC-L 126
Db 81 FCFE-GSPANRKTIGVCTAKGAPCIFIQGYVYRSGESFQSSCKYCTCLDGAVGCMPLCSM 139
QY 127 RVRRPRLMCPHRRVRSIPGHCCCEQWCEDDAKRPRKTAAPDGTAFDAVGEVEANHR--- 182
Db 140 DVLRLSPDCPPRRVKLPKCCCEWVDE-----PKDQ---TWVGPALAAVRLDPT 187
QY 183 -----NCIATYSPWSPCSTSCGLSTRTISNNACWPEOERLCLNLRPCVDIH 232
Db 188 FGDPPTMIRANCLVQTTWMSACSTCGMISTRTVNDNASCRLEKOSRLCMVRCEADLE 247
QY 233 TLKAGKCLAVYQPEASMTFLAGCISTRSYOPKCYGVCMDNRCCIPYKSKTIDVAFQC 292
Db 248 ENIKKGGKCIKRTIRKISPIFELSGCTSMKTYRAKFGCVCTDGRCCCPHRTTILPVEFKC 307
QY 293 PDGLGFSROYLWIMINACFCNLSCRNPNDIFADL---ESYPDFS 331
Db 308 PDGEVMMKMMFIKTCACHYNCPCDNDIFESLYRKMVGMA 349

```

```

RESULT 9
PCT-US96-08140-2
; Sequence 2, Application PC/TUS9608140
; GENERAL INFORMATION:
; APPLICANT: University of South Florida
; TITLE OF INVENTION: CONNECTIVE TISSUE GROWTH FACTOR
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FISH & RICHARDSON P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: US
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/08140
; FILING DATE: 30-MAY-1996
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Halle, Ph.D., Lisa A.
; REGISTRATION NUMBER: 38,347
; REFERENCE/DOCKET NUMBER: 07414/003W01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-678-5070
; TELEFAX: 619-678-5099
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 349 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US96-08140-2

```

```

Query Match 44.9%; Score 869.5; DB 4; Length 349;
Best Local Similarity 45.6%; Pred. No. 2.6e-73;
Matches 156; Conservative 53; Mismatches 98; Indels 35; Gaps 8;

QY 12 SRP---QFCWPCPCPPSP-PRCPGLVSLTTDGCCECKMCAQULGDMCTEATCDPRL 67
Db 21 SRAVGNCGSGPCPCPEPAPRCAGVSLVDGCGCCRVCAKOLGELCTERDPCDPHKL 80
QY 68 YCDYSGDRPYAIGVCAQVGVGCVLDGVYRNNQSFOPKCYNCCTCIDGAVGCTPLC-L 126
Db 81 FCFE-GSPANRKTIGVCTAKGAPCIFIQGYVYRSGESFQSSCKYCTCLDGAVGCMPLCSM 139
QY 127 RVRRPRLMCPHRRVRSIPGHCCCEQWCEDDAKRPRKTAAPDGTAFDAVGEVEANHR--- 182
Db 140 DVLRLSPDCPPRRVKLPKCCCEWVDE-----PKDQ---TWVGPALAAVRLDPT 187
QY 183 -----NCIATYSPWSPCSTSCGLSTRTISNNACWPEOERLCLNLRPCVDIH 232
Db 188 FGDPPTMIRANCLVQTTWMSACSTCGMISTRTVNDNASCRLEKOSRLCMVRCEADLE 247
QY 233 TLKAGKCLAVYQPEASMTFLAGCISTRSYOPKCYGVCMDNRCCIPYKSKTIDVAFQC 292
Db 248 ENIKKGGKCIKRTIRKISPIFELSGCTSMKTYRAKFGCVCTDGRCCCPHRTTILPVEFKC 307
QY 293 PDGLGFSROYLWIMINACFCNLSCRNPNDIFADL---ESYPDFS 331
Db 308 PDGEVMMKMMFIKTCACHYNCPCDNDIFESLYRKMVGMA 349

RESULT 10
US-08-468-847B-14
; Sequence 14, Application US/08468847B
; Patent No. 5780263
; GENERAL INFORMATION:

```

APPLICANT: Hastings, Gregg A. and Adams, Mark D.  
TITLE OF INVENTION: Human CCN-Like Growth Factor  
NUMBER OF SEQUENCES: 20  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLIAN,  
ADDRESSEE: CECCHI, STEWART & OLSTEIN  
STREET: 6 BECKER FARM ROAD  
CITY: ROSELAND  
STATE: NEW JERSEY  
COUNTRY: USA  
ZIP: 07068  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 INCH DISKETTE  
COMPUTER: IBM PS/2  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: WORD PERFECT 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/468,847B  
FILING DATE: 6 June 1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: MULLINS, J.G.  
REGISTRATION NUMBER: 33,073  
REFERENCE/DOCKET NUMBER: 325800-442  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 201-994-1700  
TELEFAX: 201-994-1744  
INFORMATION FOR SEQ ID NO: 14:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 348 AMINO ACIDS  
TYPE: AMINO ACID  
STRANDEDNESS:  
TOPOLOGY: LINEAR  
MOLECULE TYPE: PROTEIN  
US-08-468-847B-14

Query Match 44.4%; Score 860; DB 1; Length 348;  
Best Local Similarity 45.6%; Pred. No. 1.9e-72;  
Matches 156; Conservative 53; Mismatches 97; Indels 36; Gaps 9;  
OY 12 SRP---QFCWPCPCPPSP-PRCPPLVSLITDGCCECKKCAQOLDNCTEAACDPHRL 67  
DB 21 SRPAVGQNSGRCRCPDEBPAPGVSIV-DGCGCCRYCAQOLGELCTERDPCDPHKL 79  
OY 68 YCDYSGDRPRRYAIGCAQVGVGYLDGVRYNNGSFOPNCKRYNCTCIDGAVGCTPLC-L 126  
DB 80 FCDP-GSPANRRIGVCTAKDAPICFEGGVYRSGSFOSCKYQCTCIDGAVGCTPLCSM 138  
OY 127 RVRPRRLMCPHRRYSIPHCCEQWYCEDDARPRKTAAPRDTGARDVAGEVAMHR---- 182  
DB 139 DYRLSPDCPFPRRYKLPKCKCEWVDE-----PKDQ---TVGPAALAAVRLDPT 186  
OY 183 -----NCIATSPSPCSTSGGLGVSTRISVNAOCMPDEOSRLCNLRPCDVDT 232  
DB 187 FCPDPTMIRANCLVOTTEMSACSKTCGMGISTRVINDNFTCLEROSRLCMVRPEADLE 246  
OY 233 TLIRAGKCLAVYQPEASNFTLACISIRSYQPYCYGCMNRCICPYKSTIDVSPQC 292  
DB 247 ENIKKGKCIIRTPKIAKPYKFLSGCTSVKTYRAKFCGCTGRCCTPHRTTTLVEFEC 306  
OY 293 PDGLGFSROVLMINACFCNLSCRNDIPADL---ESYDPFS 331  
DB 307 PDGEVAKKMMFIKTCACHYNCPGNDIFESLYRKMVGDMA 348  
RESULT 11  
US-08-468-847B-15  
; Sequence 15, Application US/08468847B  
; Patent No. 5780263

GENERAL INFORMATION:  
APPLICANT: Hastings, Gregg A. and Adams, Mark D.  
TITLE OF INVENTION: Human CCN-Like Growth Factor  
NUMBER OF SEQUENCES: 20  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLIAN,  
ADDRESSEE: CECCHI, STEWART & OLSTEIN  
STREET: 6 BECKER FARM ROAD  
CITY: ROSELAND  
STATE: NEW JERSEY  
COUNTRY: USA  
ZIP: 07068  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 INCH DISKETTE  
COMPUTER: IBM PS/2  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: WORD PERFECT 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/468,847B  
FILING DATE: 6 June 1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: MULLINS, J.G.  
REGISTRATION NUMBER: 33,073  
REFERENCE/DOCKET NUMBER: 325800-442  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 201-994-1700  
TELEFAX: 201-994-1744  
INFORMATION FOR SEQ ID NO: 15:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 348 AMINO ACIDS  
TYPE: AMINO ACID  
STRANDEDNESS:  
TOPOLOGY: LINEAR  
MOLECULE TYPE: PROTEIN  
US-08-468-847B-15

Query Match 42.9%; Score 831.5; DB 1; Length 348;  
Best Local Similarity 44.2%; Pred. No. 8.7e-70;  
Matches 151; Conservative 56; Mismatches 100; Indels 35; Gaps 8;  
OY 12 SRP---QFCWPCPC-PPSPRCPPLVSLITDGCCECKKCAQOLDNCTEAACDPHRL 67  
DB 20 TRPATGQDCSACQCAAEAPHPCHPAGVSLVLDGCGCCRYCAQOLGELCTERDPCDPHKL 79  
OY 68 YCDYSGDRPRRYAIGCAQVGVGYLDGVRYNNGSFOPNCKRYNCTCIDGAVGCTPLC-L 126  
DB 80 FCDP-GSPANRRIGVCTAKDAPICFEGGVYRSGSFOSCKYQCTCIDGAVGCTPLCSM 138  
OY 127 RVRPRRLMCPHRRYSIPHCCEQWYCEDDARPRKTAAPRDTGARDVAGEVAMHR---- 182  
DB 139 DYRLSPDCPFPRRYKLPKCKCEWVDE-----PKDQ---TVGPAALAAVRLDPT 186  
OY 183 -----NCIATSPSPCSTSGGLGVSTRISVNAOCMPDEOSRLCNLRPCDVDT 232  
DB 187 FCPDPTMIRANCLVOTTEMSACSKTCGMGISTRVINDNFTCLEROSRLCMVRPEADLE 246  
OY 233 TLIRAGKCLAVYQPEASNFTLACISIRSYQPYCYGCMNRCICPYKSTIDVSPQC 292  
DB 247 ENIKKGKCIIRTPKIAKPYKFLSGCTSVKTYRAKFCGCTGRCCTPHRTTTLVEFEC 306  
OY 293 PDGLGFSROVLMINACFCNLSCRNDIPADL---ESYDPFS 331  
DB 307 PDGEIMKMMFIKTCACHYNCPGNDIFESLYRKMVGDMA 348  
RESULT 12  
US-08-468-847B-11  
; Sequence 11, Application US/08468847B







Tue Feb 27 15:15:19 2001

us-09-325-019-2.rai

Page 9

Oy 296 LGFSROVLWINACFCNLSCRNPNDIF 321  
: | : | | : : | : |  
Db 319 QIVKKPVMVIGTCTCHTNCPRKNEAF 344

Search completed: February 16, 2001, 16:13:10  
Job time: 22 sec

**THIS PAGE BLANK (USPTO)**